CHILD DEVELOPMENT

Child Development, March/April 2016, Volume 87, Number 2, Pages 414–428

Affecting Girls' Activity and Job Interests Through Play: The Moderating Roles of Personal Gender Salience and Game Characteristics

Emily F. Coyle and Lynn S. Liben
The Pennsylvania State University

Gender schema theory (GST) posits that children approach opportunities perceived as gender appropriate, avoiding those deemed gender inappropriate, in turn affecting gender-differentiated career trajectories. To test the hypothesis that children's gender salience filters (GSF—tendency to attend to gender) moderate these processes, 62 preschool girls (M = 4.5 years) were given GSF measures. Two weeks later, they played a computer game about occupations that manipulated the game-character's femininity (hyperfeminized *Barbie* vs. less feminized Playmobil *Jane*). Following game play, girls' interests in feminine activities showed an interaction of game condition and GSF: High-GSF girls showed intensified feminine activity interests only with Barbie; low-GSF girls showed no change with either character. Neither GSF nor game condition affected occupational interests. Implications for GST, individual differences, and occupational interventions are discussed.

Gender is a core aspect of human identity, and various processes have been proposed to account for the emergence, maintenance, and revision of gender-differentiated beliefs and behaviors. Studying such processes is important not only for evaluating developmental gender theories, but also for designing interventions intended to encourage children to eventually pursue educational and occupational activities unconstrained by cultural gender stereotypes (e.g., Eccles, 2014). Overviews of gender development (e.g., Blakemore, Berenbaum, & Liben, 2009) have identified three families of approaches that emphasize either direct environmental teaching and observational processes, or biological or evolutionary processes, or self-driven, constructive processes in which children play an active role in the way gendered material is encountered, encoded, and processed. The work we describe here was designed to test hypotheses derived from the third of these approaches.

This material is based on the work supported by the National Science Foundation Graduate Research Fellowship awarded to the first author under Grant DGE1255832. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We thank Lacey Hilliard, Cori Bower, Sherilyn Peterson, and Kathleen Ghio for their assistance with instrument development and data collection, and the anonymous reviewers and editor for their thoughtful comments and suggestions. We also appreciate the thoughtful contributions of Janet Swim and Dawn Witherspoon during the initiation of this project, which began as the first author's master's thesis, under the direction of the second author.

Correspondence concerning this article should be addressed to Emily F. Coyle, Department of Psychology, The Pennsylvania State University, 502 Moore Building, University Park, PA 16802. Electronic mail may be sent to coyle@psu.edu.

Most broadly, our work falls under the umbrella of the constructivist gender schema theory (GST) developed by Martin and Halverson (1981). A core component of GST is its emphasis on children's differential engagement with objects and activities depending on whether a particular object or activity is judged to be self-relevant. In this model, judgments about self-relevance depend, first, on children's knowledge about what is culturally defined as "for girls" versus "for boys" and, second, on their self-identification as girls or boys.

A rich empirical literature has provided support for GST by showing that at the group level, boys and girls tend to approach environmental objects, activities, and peers that are culturally defined as appropriate for their gender and tend to avoid those that are culturally defined as appropriate for the other gender (e.g., Cherney, Kelly-Vance, Gill Glover, Ruane, & Oliver Ryalls, 2003; Maccoby & Jacklin, 1974; Martin, Fabes, Hanish, & Hollenstein, 2005; Martin et al., 2012). Also supporting GST are empirical studies demonstrating that children play more with novel toys that are explicitly labeled as appropriate for their own rather than for the other gender (e.g., Cherney & Dempsey, 2010; Martin, Eisenbud, & Rose, 1995; Weisgram, Fulcher, & Dinella, 2014).

There is also strong evidence that by preschool, and perhaps even by infancy, children in the

DOI: 10.1111/cdev.12463

United States are highly knowledgeable about what is culturally defined as masculine versus feminine, including toys and occupations (e.g., Fulcher, Sutfin, & Patterson, 2008; Helwig, 1998; Jadva, Hines, & Golombok, 2010; LoBue & DeLoache, 2011). Even once knowledge of cultural gender stereotypes is well established and pervasive, however, there remains considerable variability with respect to how fully individuals believe that others should adhere to cultural gender expectations and about how strongly they endorse gender-traditional beliefs and roles for themselves (e.g., Leinbach, Hort, & Fagot, 1997; Liben & Bigler, 2002; Serbin, Poulin-Dubois, Colburne, Sen, & Eichstedt, 2001; Signorella, Bigler, & Liben, 1993; Weinraub et al., 1984). These individual differences have been shown to affect the way in which children process gender-nontraditional material. For example, the well-documented advantage for remembering stereotype-consistent better than stereotype-inconsistent gender stimuli is weaker or even absent among children who show low personal endorsement of those stereotypes (Signorella & Liben, 1984). In addition to individual differences in the strength of endorsement of cultural gender stereotypes, there are also individual differences in the extent to which individuals tend to view the world through gendered lenses, a difference Bem (1981) described as the contrast between "gender schematic" and "gender aschematic" individuals.

Individual differences in both types have been incorporated into a more recent constructive approach to gender development by Liben and Bigler (2002) that fine-tunes GST in two major ways. First, this model specifies the importance of two pathways by which children's gender differentiation evolves, leading the model to be referred to as a dual-pathway model (DPM; see Blakemore et al., 2009). One of these—the *attitudinal pathway*—is like the one proposed in GST insofar as it holds that a child's beliefs about what is culturally acceptable for girls versus boys drives the child's own involvement and actions. The other—the *personal pathway*—identifies an inverse route in which the child's personal interests and routine attention to gender are thought to drive the further development and modification of gender attitudes.

Entailed in the operation of these pathways is thus a second way in which DPM extends the earlier GST, namely, its more explicit attention to the role of individual differences. These include children's personal interests (e.g., being particularly attracted to a particular toy completely aside from whether it is linked to masculine or feminine cultural roles) and the strength of children's routine attentiveness to gender. In DPM, the latter is referred to as the child's gender salience filter (GSF), that is, the degree to which children have, readily activate, and use gender schemata in their intercourse with the world. A child who has a strong GSF is one who tends to connect observations and interactions to gender, whether or not there is any external encouragement or reason to do so.

Thus, DPM specifies the moderating role of both personal interest and gender salience in gender schema development (Liben & Bigler, 2002). It posits that a child may approach a potentially schema-modifying encounter (i.e., an encounter that would normally be avoided because of its inconsistency with cultural gender expectations) if the situation is of personal interest. Furthermore, it predicts that the degree to which the child's resulting experiences will, in turn, affect the child's gender schemata will be moderated by the strength of that child's GSF. Children with a strong GSF would be more likely to modify their own gender schemata in light of their experience. For example, a girl who happens to find trucks appealing and therefore plays with trucks would be more likely to use her personal involvement to modify her gendered beliefs about truck play if she has a strong GSF. She would be likely, for example, to modify her initial belief that "trucks are for boys" to the new belief that "trucks are for boys *and* girls." Absent the tendency to process experiences in terms of gender (i.e., a low GSF), a girl would have little reason to translate her personal interest and experience with trucks to her gendered beliefs more generally. Her truck play would likely increase her knowledge in other ways (e.g., increasing her understanding of the effects of friction from observing the truck lose speed over distance), but would not be likely to influence her beliefs about the gendered nature of truck play.

The major goal of the research reported here was to provide a test of the DPM-derived hypothesis that individual differences in GSF do, indeed, moderate children's responses to experiences that have the potential to affect the child's own gender schemata and behaviors. The particular gender-related domain we chose for our work was motivated by the applied problem identified in the introductory paragraph—attracting children to pursue occupational careers that are traditionally associated with the other sex. More specifically, we chose to focus on the goal of attracting girls to jobs traditionally stereotyped as masculine. This is an especially important goal because it is relevant to attracting women to STEM occupations (science, technology, engineering, and mathematics) that have been identified as needing more U.S. talent (e.g., Ceci & Williams, 2010; National Science Foundation, 2014) and that continue to show underrepresentation of women, particularly in engineering and technology (e.g., Ceci & Williams, 2010; Halpern et al., 2008; Hill, Corbett, & St. Rose, 2010; Liben & Coyle, 2014).

Although many kinds of gender-targeted interventions to increase women's participation in STEM have been attempted, and positive outcomes have been reported (Leaper, 2015; Liben & Coyle, 2014), there remains much room for improvement. Given that weaker or less resilient interests in some STEM fields appear to reduce girls' entry into or persistence in traditionally masculine careers (e.g., Ceci & Williams, 2010; Eccles, 2014), it may be particularly important to target interventions for children when they are very young and their ideas about gender and their specialized interests are first developing.

For this young age group, play is an important arena for developing and expressing skills and interests. For example, research shows that some modes of play (e.g., block play) promote spatial skill development (e.g., Verdine et al., 2014), a skill set associated with success in STEM (Wai, Lubinski, & Benbow, 2009). As already illustrated in the discussion of both GST and DPM, there is considerable evidence that children apply their developing understanding of gender to play, approaching what they perceive as culturally appropriate for their own gender, and avoiding what they perceive as appropriate for the other gender. Given that children tend to take on characteristics or behaviors of same-sex models (Bussey & Bandura, 1999) and are generally attracted to own-sex-typed toys (e.g., Cherney et al., 2003), high-GSF girls might be especially attracted to play with a character that is strongly feminized. From a modeling perspective, high-GSF girls might be expected to be especially drawn to masculine activities that have been modeled by a strongly feminized character rather than by a weakly feminized character. In the current study, we tested whether individual differences in GSF would have a significant impact on the consequences of play with a computer game about gender-stereotyped occupations. Specifically, we examined the role of GSF as a moderator of the effects of an intervention that exposed preschool girls to various occupations being enacted by a female game character. Onethird of the depicted occupations were stereotypically feminine jobs that were already likely to be of interest to young girls. Another third of the depicted occupations were culturally masculine jobs that were likely to be of less interest to young girls. The remaining occupations were entirely novel (Liben, Bigler, & Krogh, 2001), included to examine effects of game condition and GSF on girls' interests aside from preexisting gender stereotypes related to these jobs. A central research question was whether girls' interests in masculine jobs would be greater when the game character who enacted them was highly feminized (Barbie) rather than less feminized (a female Playmobil doll, referred to as *Jane*), and whether these patterns would vary in relation to girls' individually assessed GSFs.

Given that girls in general are more likely to be attracted to Barbie than to other dolls (Steinberg, 2004), girls in the Barbie condition were expected to be more interested in the game-depicted jobs of any kind than were girls in the Jane condition. On the basis of the personal pathway of DPM (Liben & Bigler, 2002), we also predicted that individual differences in GSF would interact with game condition: Girls with a stronger or high GSF (H-GSF) were expected to be especially strongly influenced by game play with Barbie than with Jane, whereas girls with a weaker or low GSF (L-GSF) were expected to be less susceptible to distinctions between the two game characters. Because exposure to jobs in our intervention was through the traditionally feminine medium of doll play (rather than, say, through didactic lessons about occupations as used in other work; e.g., Bigler & Liben, 1990; Weisgram & Bigler, 2006) and given the importance of some modes of play to skill development that may ultimately be useful for some jobs (Verdine et al., 2014; Wai et al., 2009), we also expected that game play would affect girls' own activity interests. We predicted that, overall, girls in the Barbie condition would show a greater increase in their interest in feminine (but not masculine) play activities with Barbie than Jane because Barbie would be more likely to prime or encourage culturally feminine behaviors, and that the pattern would again be affected by individual differences in GSF such that H-GSF girls would show a stronger effect of game-character condition than would L-GSF girls.

Before turning to a description of the methods used to address these questions, it is important to note that although Liben and Bigler (2002) assigned an important role to the GSF within their model, they offered only a general description of the construct, and neither they—nor, to our knowledge, subsequent researchers—provided an assessment tool for its measurement. To operationalize the construct here, we thus selected three gender-related constructs and measures already in the literature that are conceptually relevant to the broader construct of GSF. The first, affiliative preference for play partners of one gender, taps the personal importance of gender for interaction partners (see Serbin & Sprafkin, 1986). Children who show no particular gender-based affiliative preference are arguably less attentive to gender more generally. The second, implicit memory for gender, taps the degree to which children encode information about gender in the environment even in contexts in which there is no explicit need to do so (see Liben & Hilliard, 2010). The third, gender typicality, refers to the extent to which children perceive themselves to be similar to others of their own gender (see Carver, Yunger, & Perry, 2003; Martin, Fabes, Hanish, Leonard, & Dinella, 2011). Children who monitor and match themselves to one rather than the other gender must necessarily be paying attention to gender-linked traits and behaviors. A fuller description of each of the three constructs, assessment tools, and data supporting the integration of the three measures into a composite measure of GSF is provided in the Method section next.

Method

Participants

Before embarking on data collection, we conducted an a priori power analysis to determine the sample size. Estimates for effect size were based on previous work by Coyle and Fulcher (2011) that used similarly feminized Barbie dolls and many of the same measures used in the present study to teach preschool girls about jobs and examine effects on job self-efficacy. Data showed a very large effect size, Cohen's $d = 1.34$, equivalent to partial $\eta^2 = .31$. We used the more conservative estimate for large effects of partial $\eta^2 = .14$. Using GPower (Faul, Erdfelder, Lang, & Buchner, 2007), we estimated that to obtain 80% power with our design (2 conditions, 1 covariate, 2 additional predictors, and all interactions), we would need approximately 60 subjects.

Thus, our final sample included 62 girls who ranged in age from 48 to 78 months ($M = 53.7$ months, $SD = 7.5$ months) and were attending suburban, university, or community preschools in the Northwest (Washington state) or Northeast (Pennsylvania) United States. The schools serve primarily well-educated, middle-class families, about 75% White, 20% Asian American, and 5% African American. Children were recruited following each school's recruitment policies. For all schools, we sent letters to parents describing the study. In one school, parents were asked to return signed consent forms if they wished their child to participate; in the remaining schools, parents were asked to return a form if they preferred that their child not to participate. Approximately 75% of parents of eligible girls in participating schools agreed to their child's participation. Children were also asked if they were willing to participate, and 100% assented positively. At all schools, children were tested in a quiet room adjacent to their classroom. Data were collected between fall 2011 and spring 2012.

Procedural Overview

Girls participated in two sessions, each administered by a different experimenter to reduce the chance that children would link the two sets of activities. In Session 1, children were first given existing measures of personal interests in masculine and feminine activities using the activities subscale of the Preschool Occupations, Activities, and Traits–Personal Measure (POAT–PM; Liben & Bigler, 2002), and then given an assessment of GSF. Session 2 took place 1–2 weeks later. The second tester played with girls using a computer game that described jobs and the tools needed to perform those jobs. Following completion of the game, girls rated their interest in each of the depicted game jobs, next completed a measure of their interests in masculine and feminine occupations that had not been depicted in the game (modified POAT–PM), and then again completed the measure of personal interest in masculine and feminine activities that had been used at pretest (POAT–PM; Liben & Bigler, 2002). Both sessions ended by girls selecting stickers as a thank-you gift for participation. Next we describe in more detail the computer game, the GSF measure, and all activity and job interest measures.

Computer Game

The game involved learning about occupations enacted on a computer screen by a female character. Varied randomly between participants was whether the female character appearing in the game was hyperfeminized versus a more natural female character. The hyperfeminized character was Barbie, a well-known doll that is popular with young girls, appearing in many contexts including fashion play, computer games, and movies (Steinberg, 2004). The less feminized character (Jane) was based on a Playmobil doll. Jane shared skin tone, hair color, and hair length with Barbie, and was readily interpreted as female but had neither the narrow waist nor exaggerated breasts of Barbie.

The dolls were depicted in 12 different jobs, presented in one randomized order. Included were four traditionally feminine jobs (librarian, teacher, nurse, and florist), four traditionally masculine jobs (explorer, astronaut, firefighter, and chemist), and four novel jobs (benster, hegist, silter, and tenic, drawn from Liben et al., 2001). The outfits for each job were matched in Barbie and Jane conditions. The two versions of the game were written using the software GameMaker for Mac (YoYoGames, 2012). A computer game (rather than physical dolls) was used in order to control length of play with the characters and allow equivalent depiction of careers and settings across two different dolls that in physical form vary in size and wardrobe.

The experimenter explained that they would play a computer game to learn about different jobs and the tools needed to perform those jobs. For each job, the character appeared on the screen with three tools as illustrated in Figure 1. The experimenter explained the nature of that job and asked the child to identify which of the three tools were relevant for the job and which were irrelevant. When clicked, relevant tools jumped to the bottom of the screen and the irrelevant tool disappeared. For example, for the astronaut depicted in Figure 1, the child was told, "Barbie [Jane] is an astronaut. An astronaut is someone who flies to outer space in spaceships to learn about planets and stars." The child was then asked, "What tools does Barbie [Jane] need to help her do her job? What tool will not help her do her job?" After playing the entire game, the experimenter went back through the depicted jobs, showing the child the jobs on screen again, repeating the job description, and asking the child to rate interest in each job using an interest measure described next.

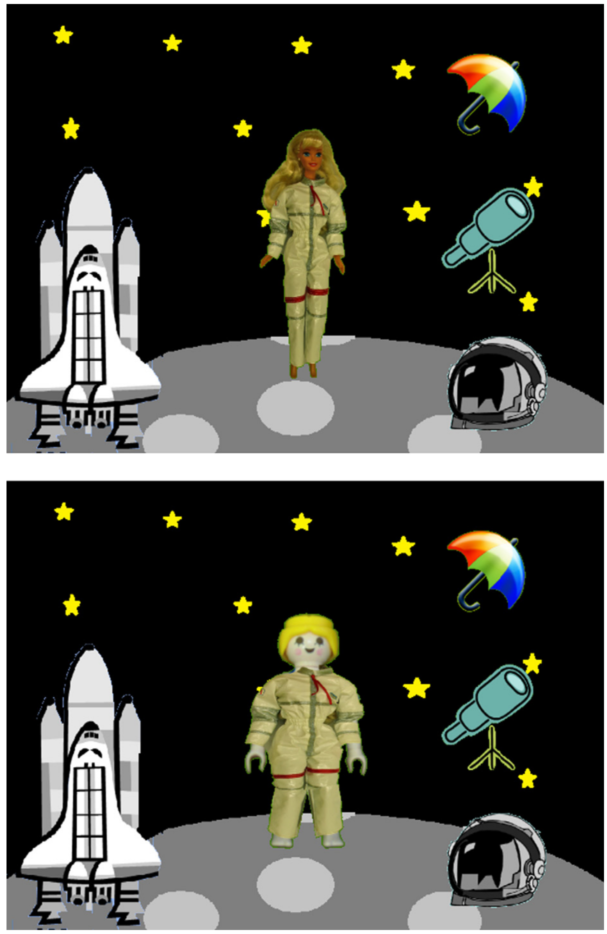

Figure 1. Sample screens for astronaut in the Barbie condition (top) and Jane (Playmobil) condition (bottom). The job-relevant items are the telescope and the helmet, and the irrelevant item is the umbrella.

GSF Measure

As explained briefly earlier, we developed a measure of GSF by drawing on existing assessments of three constructs that we identified as tapping one's attentiveness to gender—the impact of gender in making affiliative preferences, incidental memory for gender in a context in which gender was neither required nor primed by the task at hand, and girls' sense of their own gender typicality. We administered one existing assessment for each of these component constructs, and then factor analyzed responses to determine whether the three sets of items would yield a cohesive measure of the GSF construct.

Serbin and Sprafkin's (1986) measure of gendered affiliation was used to measure the degree to which gender is a salient category for children for affiliative preference. Children were first given a deck of cards showing five pairs of photos, each with a man and a woman with neutral expressions and no props, and asked with whom they would prefer to play. Based on their choices to the first deck, children were given one of two second decks. Both second decks again contained five cards, each with a man and woman. If the child selected mostly women in the first deck, the pairs in the second deck showed women with blank expressions and men with smiling expressions and holding an interesting game or toy. If the child selected mostly men in the first deck, the second deck showed the inverse parings, that is, men with blank expressions and no props, and women smiling and holding a toy. The affiliation score was the proportion of selections in the second deck that matched the prominent gender preference in the first deck. A high score indicates that the child chooses play partners based on preferred gender, even when that choice means rejecting a friendlier and better equipped play partner of the nonpreferred gender. A low score indicates that the child uses criteria other than gender for choosing play partners. This scale has shown high internal validity and test–retest reliability in scale development and previous use (Serbin & Sprafkin, 1986) and its reliability was also acceptable in the present study ($\alpha = .74$).

The individual difference measure of incidental memory for gender was drawn from a method previously developed by Liben and Hilliard (2010) to measure children's gender vigilance in different classroom contexts. The task taps the degree to which children encode information about gender in the course of making judgments about the appeal of 10 gender-neutral toys. Children were first shown 10 cards, each containing a photograph of a toy being played with by a girl (5 cards) or a boy (5 cards). Cards were presented in a single randomly determined order with specific toy–gender pairings counterbalanced across participants. Children were first asked to rate how much they would like to play with each depicted toy. After a delay, children were presented with 10 new cards showing just the toys (i.e., photographed without a child). For each card, participants were asked whether the child originally seen playing with the toy was a boy or a girl. The proportion of items for which gender was recalled correctly served as a measure of children's attention to gender. Reliability was adequate ($\alpha = .64$).

The measure of gender typicality was adapted for preschool children from a measure that had been designed by Patterson (2012) to assess the degree to which school-age girls perceived themselves as being similar to other girls. Girls were asked to rate their own similarity to girls in terms of appearance, interest, competencies, and behaviors on a 3-point scale for which response options were *not at all*, *some*, or *a lot*, accompanied by graphic representations of glasses that were completely empty, half full, or completely full. Sample items include, "Think about the things that most girls do. How much are you like that?" and "Think about the things that girls are good at. How much are you like that?" Parallel filler items asked about similarity to boys, but these responses were not used here. Responses were averaged to yield scores between 1 and 3 indicating self-perceived gender typicality. Cronbach's alpha showed adequate reliability (.70).

To examine how the three assessments of GSF (gendered affiliation, gender memory, and gender typicality) were interrelated, we conducted a factor analysis. Although conceptually all three measures should be components of the GSF, we used a conservative exploratory (rather than confirmatory) approach that permitted the analysis to reveal one, two, or three factors (rather than forcing the measures to represent a single latent construct). Using principal axis factoring without rotation, we found all three components loaded on a single factor that explained 38.10% of the variance. The scree plot confirmed a single factor solution was preferable given the steep decline in eigenvalues after the single-factor solution. The factor matrix and communalities are presented in Table 1. Because the three measures loaded on a single factor, a single GSF composite score was calculated that summed standard scores of the gendered affiliation, gender memory, and gender typicality measures.

Interest Measures

Activity Interests

To assess the degree to which children are interested in culturally feminine and masculine activities, children were given the activity subscale of the PM of the POAT sex-typing scale (Liben & Bigler, 2002; Liben, Bigler, Shechner, & Arthur, 2006). The POAT–PM contains six masculine, six feminine, and two filler neutral items, presented in one randomly determined order. Children are asked to rate their interest in each activity using a 3-point Likert scale. Response options are *not at all*, *some*, and *a*

Table 1
Exploratory Factor Analysis of Gendered Affiliation, Gender Memory, and Gender Typicality

	Factor 1 (GSF)	Communality
Gendered affiliative preference	0.468	0.219
Gender memory	0.222	0.049
Gender typicality	0.152	0.023

Note. Principal axis factoring was used. Factor 1 explained 38.10% of the variance. GSF = gender salience filter.

lot, including accompanying pictures of glasses as described for the gender typicality measure. Responses are scored, respectively, as 0, 1, and 2; responses to masculine and feminine items are averaged separately to yield masculine and feminine activity interest scores. An illustrative item (feminine example) shows a picture of a baby doll and asks, "How much do you like to play with baby dolls?" In the current study, the activity scale was given during Session 1 and in Session 2 after the computer game had ended. Prior work has shown good reliability for the measure, with Cronbach's alphas ranging between .70 and .90 (e.g., Blakemore & Phillips, 2014; Friedman, Leaper, & Bigler, 2007; Hilliard & Liben, 2010; Shechner & Lobel, 2010). Reliability in the present study was similar, with Cronbach's alphas ranging over sessions and groups between .66 and .84.

Job Interests

Parallel measures were used to assess children's interests in culturally feminine and masculine jobs, although to avoid the possibility that job-related measures would contaminate game play, job interest measures were administered only in Session 2 after game play had been completed. The first job-interest measure addressed jobs that had been included in the computer game. Specifically, after game play was completed, the experimenter displayed each job one at a time on the computer screen, again, naming and defining the job (as in the original game). For each, the experimenter asked the child to rate her interest in the job using the same 3-point response options described earlier. Because ratings were made while screens were visible, and because screens showed the character in appropriate work settings and with correct job-related tools, the interest rating assessment did not make demands on children's memories for the depicted jobs.

As a second job-interest measure, children were given a modified version of the occupation subscale of the POAT (POAT–PM), which, like the activity measure, contains six masculine, six feminine, and two filler neutral jobs, presented in one randomly determined order. A sample (masculine) item from the POAT–PM shows a picture of a car and repair tools and asks the child, "This job is a car mechanic, someone who fixes people's cars. How much would you like to be a car mechanic?" The scale was modified only insofar as three jobs that had been used in the computer game (astronaut, florist, and nurse) were dropped and three other jobs (garbage collector, babysitter, and house cleaner) were added. The substitute items were selected from a longer (34-item) version of the scale that had been given to other samples of preschool children as part of scale development (Liben et al., 2006). Items selected were jobs that had received similar ratings to the dropped items in samples given the long version of the scale.

Prior work has demonstrated that reliabilities for the occupation POAT–PM subscale have been satisfactory, with Cronbach's alphas ranging between .78 and .86 (e.g., Blakemore & Phillips, 2014; Hilliard & Liben, 2010; Shechner & Lobel, 2010). In the current study, reliabilities for the occupation subscale of the POAT–PM were good for masculine jobs (α = .72), but low for feminine jobs (α = .50). A similar pattern was evident in reliabilities on the game-job interest measure: For the masculine and novel game jobs, Cronbach's alphas were .67 and .61, respectively, but for the feminine jobs, alpha was .37. Given that the job-interest measures were administered directly after children had played with the job-related game, we infer that the unusually low reliabilities for feminine job interests in the current study are likely to be the result of children's immediatelyprior participation in the game intervention, an issue we consider in more detail in the Discussion section.

Results

Preliminary Analysis

To determine the need for covariates, we first examined correlations among demographic variables (age, race, school, and geographic location) and key study variables. These revealed no significant associations, and thus, demographic covariates were not included in later analyses. A correlation matrix is presented in Table 2. To test comparability of the two game-condition groups that had been formed by random assignment, we compared the two groups' initial activity interests (assessed by the activities subscale of the POAT–PM given in Session 1). These analyses revealed no significant differences between the two groups' interests in either feminine or masculine activities, $ts(60)$ = 0.40 and 0.53, ps = .694 and .599, respectively. Descriptive statistics for variables of interest are presented in Table 3.

Overview of Analyses

As explained in the Introduction, we hypothesized that the hyperfeminized game character (Barbie) would be more effective in influencing

Table 2
Correlations Among Key Variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Composite GSF score (S1)	—											
2. Gendered affiliation (S1)	.64**	—										
3. Gender memory (S1)	.61**	.11	—									
4. Gender typicality (S1)	.60**	.07	.03	—								
5. Game F jobs	.09	−.12	.08	.22	—							
6. Game N jobs	.06	−.10	.08	.12	.44**	—						
7. Game M jobs	−.05	−.15	−.04	.09	.54**	.52**	—					
8. POAT F jobs	.03	−.11	.02	.15	.47**	.39**	.43**	—				
9. POAT M jobs	−.08	−.18	.06	−.03	.33**	.39**	.56**	.13	—			
10. POAT F activities (S1)	.12	.07	−.13	.27*	.21	.17	.13	.35**	−.03	—		
11. POAT F activities	.27*	.17	−.06	.39**	.38**	.33**	.22	.40**	.04	.71**	—	
12. POAT M activities (S1)	−.26*	−.40**	−.05	−.03	.23	.43**	.46**	.15	.56**	.21	.10	—
13. POAT M activities	−.23	−.23	−.15	−.04	.32*	.40**	.50**	.18	.50**	.06	.08	.66**

Note. S1 indicates measures given in the first session. All other measures were given in the second session after playing the game. Variables 2–4 are subcomponents of the gender salience filter (GSF) score (Variable 1). Variables 5–7 refer to girls' personal interest in game-depicted feminine (F), novel (N), and masculine (M) jobs; Variables 8–13 are girls' interests in F or M POAT jobs and activities. POAT = Preschool Occupations, Activities, and Traits.
*$p < .05$. **$p < .01$.

Table 3
Descriptive Statistics for All Measures

Measure	Min	Max	*M*	*SD*
Gender salience measures				
Composite GSF score	−5.41	3.20	0.00	1.85
Gendered affiliation (proportion; S1)	0.00	1.00	0.59	0.32
Gender memory (proportion; S1)	0.00	1.00	0.81	0.19
Gender typicality (S1)	1.60	3.00	2.61	0.42
Job interests				
Game-depicted feminine jobs	1.25	3.00	2.22	0.48
Game-depicted novel jobs	1.00	3.00	2.18	0.55
Game-depicted masculine jobs	1.00	3.00	1.94	0.58
Feminine POAT-PM occupations	1.33	3.00	2.30	0.43
Masculine POAT-PM occupations	1.00	3.00	1.79	0.52
Activity interests				
Feminine POAT-PM activities (S1)	1.17	3.00	2.46	0.46
Feminine POAT-PM activities	1.00	3.00	2.43	0.59
Masculine POAT-PM activities (S1)	1.00	2.83	1.65	0.47
Masculine POAT-PM activities	1.00	3.00	1.62	0.54

Note. S1 indicates measures given in the first session. All other measures were given in the second session. GSF was the summed standard scores of gendered affiliation, gender memory, and own-sex typicality. Gendered affiliation and gender memory are proportion scores as noted above. The remaining scales potentially range from 1 to 3 and are reported with raw numbers. GSF = gender salience filter; POAT–PM = Preschool Occupations, Activities, and Traits–Personal Measure.

girls' interests than would the more typical character (Jane), an effect expected to be especially strong for H-GSF girls. We also expected to find—in accord with a large body of prior work—that even by preschool, girls would show significantly more interest in jobs and activities that are culturally viewed as feminine than those culturally viewed as masculine. Next we report analyses conducted to test these hypotheses with respect to girls' interests in (a) jobs depicted in the computer game itself, (b) jobs included in the occupational subscale of the POAT–PM, and (c) activities included in the activity subscale of the POAT–PM.

Interest in Game Jobs

To test the hypothesized main effects and interactions of character type and GSF score on interest in the jobs depicted in the computer game, we ran a series of multiple regressions in which game condition, GSF score, and the Condition × GSF interaction were entered stepwise to predict each of three dependent variables: girls' interest in depicted feminine jobs, masculine jobs, and novel jobs. None of the variables (i.e., condition, GSF score, or Condition × GSF) significantly predicted girls' interest in any of the three job types. Statistics for the three variables were, respectively, for feminine jobs: $R^2 = .02$, βs = −.06, .32, and −.26, *p*s = .639, .451, and .545; for masculine jobs: $R^2 = .01$, βs = −.01, .15, and −.24, *p*s = .961, .728, and .567; and for novel jobs: $R^2 = .04$, βs = .05, −.50, and .57, *p*s = .709, .236, and .177.

To compare interest across types of game jobs, we conducted a 3 (job type: feminine vs. masculine vs. novel) × 2 (condition: Barbie vs. Jane) × 2 (GSF level: high vs. low) mixed-model analysis of variance (ANOVA) in which job type was a within-subjects variable, and condition and GSF were between-subjects variables. A median split was used to divide girls into high- versus low-GSF groups. The analysis showed a significant main effect of job type, such that girls were significantly more interested in feminine (M = 2.22, SD = 0.48) and novel jobs (M = 2.18, SD = 0.55) than in masculine jobs (M = 1.94, SD = 0.58), $F(2, 114) = 9.28$, $p < .001$, $\eta^2 = .14$. Remaining main effects and interactions were not significant.

Interest in Nongame (POAT) Jobs

Comparable analyses were conducted to examine hypothesized main effects and interactions on girls' interests in culturally feminine and masculine jobs contained on the POAT–PM measure. Specifically, we again conducted stepwise multiple regressions in which game condition, GSF score, and Condition × GSF were entered as predictors of girls' interests in feminine and masculine occupations. None of these factors was a significant predictor of girls' interest in either feminine jobs (R^2 = .003, βs = −.04, −.03, and .07, *p*s = .755, .953, and .877, for condition, GSF score, and Condition × GSF, respectively) or masculine jobs (R^2 = .02, βs = .07, −.30, and .22, *p*s = .617, .473, and .592, respectively).

To compare girls' interest for different job types included in the POAT (i.e., masculine vs. feminine jobs), we again conducted a mixed-model ANOVA, specifically, a 2 (job type: POAT–PM feminine vs. POAT–PM masculine) × 2 (condition: Barbie vs. Jane) × 2 (GSF level: high vs. low) analysis in which job type was a within-subjects variable, and game condition and GSF were between-subjects variables. There was again a significant main effect of job type such that girls were significantly more interested in feminine (M = 2.30, SD = 0.43) than masculine (M = 1.79, SD = 0.52) jobs, $F(1, 56) = 39.27$, $p < .001$, $\eta^2 = .41$, regardless of game condition. Remaining main effects and interactions were not significant.

Interest in Activities

As explained in the Method section, although job interests were tested only after the child had played the computer game (to avoid having initial questions about job interests interfering with game play or its effects), activity interests could be assessed both before and after game play. We expected that playing a game with a hyperfeminized character would be likely to prime and perhaps exacerbate interest in traditionally feminine activities. Thus, we predicted that girls in the Barbie condition—relative to girls in the Jane condition—would show a greater increase in their interest in feminine (but not masculine) activities between the two sessions.

To test hypothesized main effects and interactions of character type and GSF level on change in POAT–PM feminine activity interests, we conducted an analysis of covariance (ANCOVA) controlling for baseline interests. Pretest feminine activity interests, GSF score, condition, Condition × GSF, Condition × Pretest Feminine Interests, GSF × Pretest Feminine Interests, and GSF × Condition × Pretest Feminine Interests were entered to predict posttest feminine interests. The model was significant, $F(7, 60) = 13.94$, $p < .001$, and accounted for 65% of the variance. The betas for pretest feminine interests and for GSF score were both significant in the context of the other predictors (βs = 1.07 and −5.78, *p*s = .001 and .003, respectively). Girls with greater pretest feminine interests and girls with a higher GSF score had greater posttest feminine interests. The beta for Condition × GSF was significant in the context of the other predictors (βs = 5.90, *p*s = .002). Girls in the Barbie condition with a stronger GSF also had greater posttest feminine interests, controlling for pretest interests. This interaction is depicted in Figure 2. The beta for GSF × Pretest Feminine Interests was also significant (β = 5.38, p = .004). Girls with a higher GSF score as well as greater pretest

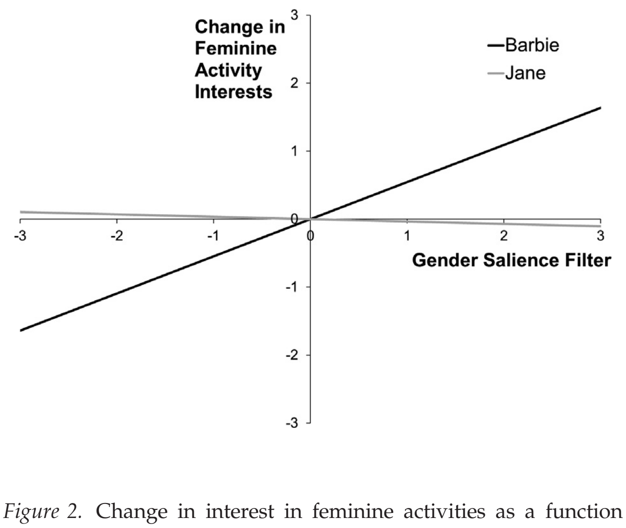

Figure 2. Change in interest in feminine activities as a function of girls' gender salience filter (GSF) scores and game condition. As explained in the text, feminine activity interests were measured using the Preschool Occupations, Activities, and Traits–Personal Measure feminine activities subscale (reported here with standard scores); GSF was the sum of three standardized scores: gendered affiliation, gender memory, and own-sex typicality.

Table 4
ANCOVA Predicting Posttest Feminine Activity Interests

Predictor	β	*t*	*p*
Pretest POAT feminine activity interests	1.07	3.58	.001
GSF	−5.78	−3.17	.003
Game condition	0.56	1.14	.258
GSF × Condition	5.90	3.33	.002
Pretest Feminine Interests × GSF	5.38	2.99	.004
Pretest Feminine Interests × Condition	−0.53	−0.98	.332
Pretest Feminine Interests × GSF × Condition	−5.33	−3.03	.004

Note. R^2 = .65. ANCOVA = analysis of covariance; GSF = gender salience filter; POAT = Preschool Occupations, Activities, and Traits.

Table 5
ANCOVA Predicting Posttest Masculine Activity Interests

Predictor	β	*t*	*p*
Pretest POAT masculine activity interests	0.38	1.10	.275
GSF	−1.13	−0.85	.402
Game condition	−0.33	−0.83	.410
GSF × Condition	1.31	0.84	.408
Pretest Masculine Interests × GSF	0.97	0.72	.476
Pretest Masculine Interests × Condition	0.43	0.86	.396
Pretest Masculine Interests × GSF × Condition	−1.17	−0.75	.457

Note. R^2 = .45. ANCOVA = analysis of covariance; GSF = gender salience filter; POAT = Preschool Occupations, Activities, and Traits.

feminine interests also had greater posttest feminine interests. Finally, the three-way interaction of GSF × Condition × Pretest Feminine Interests was significant (β = −5.33, p = .004). Examination of the plots showed that girls already at ceiling in their pretest feminine interests could not intensify in their interest from pre- to posttest. Girls with lower pretest feminine interests who played with Barbie—but not those who played with Jane—showed intensified feminine interests between pretest and posttest. Complete model statistics are presented in Table 4.

A similar ANCOVA was run to identify predictors of postgame play masculine interests. Specifically, pretest masculine activity interests, GSF score, condition, Condition × GSF, Condition × Pretest Masculine Interests, GSF × Pretest Masculine Interests, and GSF × Condition × Pretest Masculine Interests were entered as predictor variables with postgame masculine interests as the criterion variable. Although the model was significant, R^2 = .45, F(7, 60) = 6.23, p < .001, there were no individually significant betas. Girls did not increase in their interest in masculine activities from pretest to posttest, regardless of condition. Complete model statistics are presented in Table 5.

To compare activity interest across types of activities, a 2 (item type: masculine vs. feminine) × 2 (time: pre- vs. postgame) repeated measures ANOVA was conducted, with condition (Barbie vs. Jane) and GSF level (high vs. low) as between-subjects factors. There was a main effect of item type, such that all girls, regardless of GSF level or condition, were more interested in feminine (M = 2.45, SD = .53) than masculine (M = 1.64, SD = .51), activities, F(1, 57) = 109.01, p < .001, η^2 = .66. There was no significant main effect of time, F(1, 57) = 1.87, p = .176, η^2 = .03. Finally, there was a significant Item Type × GSF Level interaction, F(1, 57) = 7.17, p = .010, η^2 = .11, such that the difference between interest in feminine versus masculine activities was greater for H-GSF girls (M for feminine = 2.55, M for masculine = 1.52) than it was for L-GSF girls (M for feminine = 2.35, M for masculine = 1.74), pooled across game condition. Figure 3 depicts these findings graphically.

There was also a significant interaction among condition, time, and GSF level, F(1, 57) = 5.80, p = .019, η^2 = .09. In the Barbie condition, interest in activities (regardless of type) decreased from pre- to posttest for L-GSF girls (M pretest = 2.10, SE = 0.10, M posttest = 1.91, SE = 0.12), but

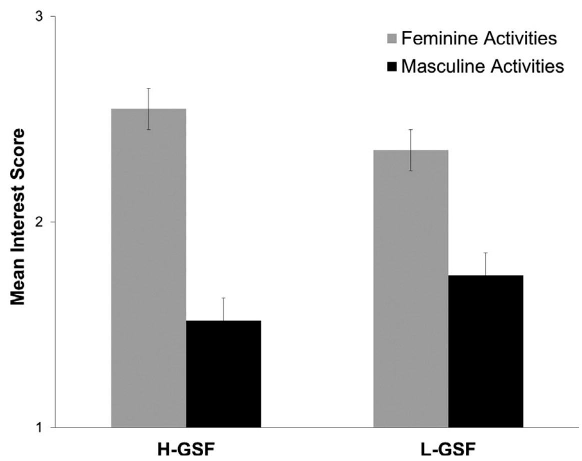

Figure 3. Interest in masculine and feminine activities for high- and low gender salience filter girls (H-GSF vs. L-GSF). Interest in masculine and feminine activities was measured by the Preschool Occupations, Activities, and Traits–Personal Measure activities subscale (rated on a 3-point Likert scale of 1 [*not at all*], 2 [*some*], and 3 [*a lot*]). GSF level was determined by a median split on the GSF score (sum of three standardized scores: gendered affiliation, gender memory, and girl typicality). Error bars represent standard error.

increased from pre- to posttest for H-GSF girls (*M* pretest = 1.99, *SE* = 0.09, *M* posttest = 2.12, *SE* = 0.10). In the Jane condition, interest in activities showed little or no change, regardless of GSF level.

Discussion

As noted in the Introduction to this article, varied theoretical processes have been hypothesized to account for gender-differentiated development during childhood. Evaluating the impact of these hypothetical processes is important not only for testing alternative theoretical models, but also for informing the practical goals of designing first, environments that minimize gender-based constraints on children's educational and occupational pursuits, and second, intervention programs that remediate unwanted consequences of gender constraints that have already taken their toll. These practical goals are of particular importance in the domain of STEM because there is a long history of research showing that, compared to boys and men, girls and women are less likely to pursue or persist in STEM careers, particularly those in technology and engineering (e.g., Ceci & Williams, 2010; Eccles, 2014; Ericson, 2014; Liben & Coyle, 2014; National Science Foundation, 2013; Petersen & Hyde, 2014).

In broad stroke, the theoretical perspective within which the current research is grounded is gender constructivism, an approach that holds that children's own qualities affect their engagement with opportunities that are potentially available in the environment. The more specific construct examined here is the GSF posited within the dual pathways model proposed by Liben and Bigler (2002). What is especially important about GSF in this model is that it points to the impact of individual differences within the operation of more universal constructivist processes. Thus, although environmental theories may look largely to the surrounding context to explain or remediate developmental pathways (e.g., focusing on how social environments differentially reinforce or punish culturally feminine or masculine behaviors in boys and girls), constructivist theories also look to qualities of individual children as potential factors in explaining or intervening in developmental outcomes.

To the best of our knowledge, this is the first empirical study to test the hypothesized relevance of a child's GSF on the impact of a given educational intervention. The particular intervention—exposure to a hyperfeminized model enacting traditionally masculine jobs—is one that parallels existing efforts aimed at increasing the appeal of traditionally masculine jobs for girls and women. Illustrative is the *Science Cheerleader* program (Cavalier, 2014), which enlists highly attractive women who were previously professional cheerleaders for the National Football League or National Basketball Association and who are currently employed in STEM fields. The group travels to different events (e.g., fairs, sports events, schools) and gives cheerleading performances in which they execute routines that include cheers for science. The rationale appears to be that exposing children to highly feminized models who describe their STEM careers and STEM activities will lead girls to see such careers as compatible with, rather than antithetical to, feminine qualities.

Similar to such intervention approaches, the Barbie condition of the computer game used in the current study shows girls a hyperfeminized character—Barbie—enacting various jobs, including some that are traditionally masculine. Of interest was whether playing this game would change participants' interests in gendered domains, and whether the impact of the game would differ in relation to girls' GSFs.

With respect to activities, results were consistent with the hypothesized importance of girls' own qualities in interaction with game condition. Most strikingly, the effect of playing the computer game in the Barbie condition varied in relation to GSF: H-GSF girls playing with Barbie showed intensified interests in feminine activities after game play, whereas L-GSF girls showed no such effect. Play with the more typically feminine doll character, Jane, had no demonstrable impact on girls' activity interests irrespective of individual differences in girls' GSF.

Consistent with previous theories and empirical research (e.g., Blakemore et al., 2009; Maccoby & Jacklin, 1974; Martin & Halverson, 1981), girls in the present study preferred feminine to masculine activities. The contrast in interest ratings for these two types of activities also varied with girls' GSF level. In comparison to L-GSF girls, H-GSF girls showed a more dramatic difference in relative preference for feminine versus masculine activities (pooling across pre- and posttest). In addition, the participant groups differed with respect to the way in which their ratings changed as a result of playing the game. For H-GSF girls in the Barbie condition, activity interest increased from pre- to posttest. This effect appears to have been driven largely by responses to feminine activities, although limited statistical power prevents us from testing the implied four-way interaction among GSF × Condition × Time × Activity Type (masculine vs. feminine). The L-GSF girls in the Barbie condition instead showed a decreased level of interest in activities between pre- and posttest. The difference in direction of the effect for H-GSF versus L-GSF girls playing with Barbie suggests the differential impact of a highly feminized character on each group. It appears that for H-GSF girls, Barbie was appealing and therefore increased activity interests, whereas for L-GSF girls, Barbie was unappealing, therefore decreasing activity interests. In the Jane condition, there was little change in interest ratings between pre- and posttests for either H-GSF or L-GSF girls.

Also of relevance from the perspective of the practical goal of increasing girls' participation in STEM careers are findings on girls' occupational interests. Again as anticipated, the girls in this study—irrespective of the strength of GSF and irrespective of game condition—expressed significantly less interest in culturally masculine occupations enacted in the game than they did in jobs from either of the other two occupational categories, that is, feminine or novel occupations. Importantly, as described earlier, although game condition interacted with GSF to increase girls' personal interests in traditionally feminine activities, there was no evidence that either game character or GSF moderated the impact of game play on girls' interests in masculine occupations. This was true for data concerning interest in jobs that were included in the computer game as well as interest in jobs that were queried as part of the POAT–PM scale. Given that the reliabilities were high for masculine game-depicted jobs as well as for the masculine POAT–PM jobs (even when reliabilities were calculated separately for each of the numerically small [n = ~15] game Condition × GSF cells), we conclude that the measures worked as intended and that the intervention truly failed to increase interest in masculine jobs. Thus, the data from the current study offer no support for the idea that girls—even those with a high GSF—exposed to a hyperfeminized character enacting traditionally masculine jobs will be drawn to those jobs, or to other traditionally masculine jobs that had not been included within the game itself. Consistent with DPM, we interpret the failure of Barbie to increase even H-GSF girls' interest in masculine jobs as consistent with GST insofar as it suggests that the girls probably dismissed masculine jobs out of hand as "not for me," and thus they had no impetus to consider or process them further.

Why might the game intervention have had so little impact on girls' interest in masculine occupations? One possibility might be that children had not fully processed the counterstereotyped job portrayals (e.g., a female astronaut). Indeed, previous research has demonstrated that children have difficulty processing gender-nontraditional information, often distorting it to transform it into a gender traditional stimulus, or forgetting it entirely (e.g., Koblinsky, Cruse, & Sugawara, 1978; Liben & Signorella, 1980). However, the finding in the current study that girls (regardless of GSF level or condition) gave significantly higher interest ratings to feminine than to masculine jobs suggests that participants did, in fact, process the gendered nature of the jobs correctly. Thus, memory distortions are unlikely to account for the intervention findings.

Other research has demonstrated the tenacity of children's own occupational interests in the face of short interventions, even when those interventions showed significant effects in other arenas. For example, Bigler and Liben (1990) gave children classroom lessons about jobs as part of a career education unit. Lessons for the experimental group (but not for the control group) included instruction about the irrelevance of gender for entering occupations. Results showed that children in the experimental group later had significantly lower gender stereotypes and significantly better memory for counterstereotypic material such as a story about a female dentist. Yet, these children showed no greater personal interest in gender-nontraditional jobs than did peers in the control group. As Bigler and Liben noted, occupational interests may be developed over years, and preschoolers may already have reasonably well-established ideas about their interests by the time they experienced the intervention, making it difficult to change their interests quickly. Similarly, a short, single-session intervention like the one used in the current study may have been insufficient to modify girls' established preferences for feminine jobs over masculine jobs. Future research is needed to evaluate whether a game that offered more intense exposure to masculine jobs (e.g., by a higher proportion of masculine jobs or by additional play sessions) would achieve more success.

As argued in DPM, it is also important to consider the role of personal interests that lie outside gender per se (i.e., interests that are related to neither GSF nor gender attitudes). As explained briefly earlier, and in more detail by Liben and Bigler (2002), DPM posits that a child's decision to approach an object or event available in the environment depends not only on whether that object or event is viewed as gender appropriate, but also on whether it happens to engage the child's personal interests or talents. The data in the present study showed that girls' self-reported interests in feminine jobs (both those in the game and those contained in the POAT scale) were quite variable. Because jobs falling within the feminine category are seen as gender acceptable, for these jobs, girls may determine their own interest in each job by considering the fit between (a) the individual job's perceived activities and demands and (b) their own particular profile of likes, dislikes, and talents. Under such circumstances, judgments would vary within the group of feminine jobs category, a pattern that is consistent with the response variability (and thus atypically low within-category reliability) observed on the feminine portion of the POAT–PM reported earlier.

Before closing, it is important to note that the present study focused exclusively on girls and used a particular type of intervention because of our interest in examining the role of GSF in a context with important societal implications—the gender gap in STEM—and with a type of experience that is actually being used in existing interventions (e.g., *Science Cheerleader*). It must be remembered, however, that the GSF is a general construct, hypothesized to operate in the gender development of boys as well as girls. Future work should thus seek to test the generality of these processes by testing the role of GSF in boys.

Additionally, it is important for future work to examine the operation of the processes considered here in far more diverse samples to test whether mechanisms are similar across the fuller ethnic, socioeconomic, and geographic diversity of children both within and beyond the United States. Such research could also have important implications for developing effective interventions to encourage participation in STEM from other currently underrepresented groups. There is already compelling evidence that the broader context has important implications in the STEM domain. Illustrative is research showing that the STEM gender gap varies in relation to nation-level indicators of gender equity (e.g., Else-Quest, Hyde, & Linn, 2010) and thus findings from one context cannot necessarily be generalized to another. To be successful, interventions will need to take both child-level and context-level factors into account.

In conclusion, the current study suggests that interventions that seek to capitalize on girls' attraction to hyperfeminized models in an attempt to increase interest in gender-nontraditional domains may actually lead to more—rather than to less—strongly gendered interests in some girls. These findings underscore the more general point that it is important to evaluate gender-targeted programs not only with respect to evidence of intended, desired outcomes, but also with respect to unintended and potentially unwanted consequences (see Liben & Coyle, 2014). As interventions are designed, implemented, and evaluated, it is critical to monitor not only outcomes of participant or student groups as conglomerate entities, but also to examine outcomes of individuals within those groups. As demonstrated by the findings of the current study, individuals may respond differentially to the same intervention. At the broadest level, the current study demonstrates the value of addressing constructivist, child-based processes in studying gender development. At a more specific level, this research offers a means for measuring GSF strength, and provides empirical evidence that GSF is an important individual quality to consider when designing and evaluating gender-targeted interventions intended to reduce gender gaps in educational and occupational settings.

References

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, *88*, 354–364. doi:10.1037/0033-295X.88.4.354.

Bigler, R. S., & Liben, L. S. (1990). The role of attitudes and interventions in gender-schematic processing. *Child Development*, *61*, 1440–1452. doi:10.1111/j.1467-8624.1990.tb02873.x

Blakemore, J. E. O., Berenbaum, S. A., & Liben, L. S. (2009). *Gender development*. New York, NY: Psychology Press.

Blakemore, J. E. O., & Phillips, E. L. (2014, October). *Do parents and children agree on children's gender-related interests and traits?* Poster presented at the Gender Development Research Conference, San Francisco, CA.

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, *106*, 676–713. doi:10.1037/0033-295X.106.4.676

Carver, P. R., Yunger, J. L., & Perry, D. G. (2003). Gender identity and adjustment in middle childhood. *Sex Roles*, *49*, 95–109. doi:10.1023/A:1024423012063

Cavalier, D. (2014). *Science Cheerleader*. Retrieved from www.sciencecheerleader.com

Ceci, S. J., & Williams, W. M. (2010). *The mathematics of sex: How biology and society conspire to limit talented women and girls*. New York, NY: Oxford University Press.

Cherney, I. D., & Dempsey, J. (2010). Young children's classification, stereotyping and play behaviour for gender neutral and ambiguous toys. *Educational Psychology*, *30*, 651–669. doi:10.1080/01443410.2010.498416

Cherney, I. D., Kelly-Vance, L., Gill Glover, K., Ruane, A., & Oliver Ryalls, B. (2003). The effects of stereotyped toys and gender on play assessment in children aged 18–47 months. *Educational Psychology*, *23*, 95–106. doi:10.1080/01443410303222

Coyle, E. F., & Fulcher, M. (2011, March). *Toys as models: How Barbie may impact girls' possible self content*. Poster presented at the Society for Research in Child Development, Montreal, Canada.

Eccles, J. S. (2014). Gender and achievement choices. In E. T. Gershoff, R. S. Mistry, & D. A. Crosby (Eds.), *Societal contexts of child development: Pathways of influence and implications for practice and policy* (pp. 19–34). New York, NY: Oxford University Press.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*, 103–127. doi:10.1037/a0018053

Ericson, B. (2014). *AP data for the United States*. Retrieved from http://home.cc.gatech.edu/ice-gt/321

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. doi:10.3758/BF03193146

Friedman, C. K., Leaper, C., & Bigler, R. S. (2007). Do mothers' gender-related attitudes or comments predict young children's gender beliefs? *Parenting: Science and Practice*, *7*, 357–366. doi: 10.1080/15295190701665656

Fulcher, M., Sutfin, E. L., & Patterson, C. J. (2008). Individual differences in gender development: Associations with parental sexual orientation, attitudes, and division of labor. *Sex Roles*, *58*, 330–341. doi:10.1007/s11199-007-9348-4

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2008). Sex, math and scientific achievement. *Scientific American*, *18*, 44–51. doi:10.1038/scientificamericanmind1207-44

Helwig, A. A. (1998). Gender-role stereotyping: Testing theory with a longitudinal sample. *Sex Roles*, *38*, 403–423. doi:10.1023/A:1018757821850

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. Washington, DC: American Association of University Women.

Hilliard, L. J., & Liben, L. S. (2010). Differing levels of gender salience in preschool classrooms: Effects on children's gender attitudes and intergroup bias. *Child Development*, *81*, 1787–1798. doi:10.1111/j.1467-8624.2010.01510.x

Jadva, V., Hines, M., & Golombok, S. (2010). Infants' preferences for toys, colors, and shapes: Sex differences and similarities. *Archives of Sexual Behavior*, *39*, 1261–1273. doi:10.1007/s10508-010-9618-z

Koblinsky, S. G., Cruse, D. F., & Sugawara, A. I. (1978). Sex role stereotypes and children's memory for story content. *Child Development*, *49*, 452–458. doi:10.2307/1128710

Leaper, C. (2015). Gender and social-cognitive development. In L. S. Liben & U. Müller (Eds.), *Handbook of child psychology and developmental science* (Vol. 2, 7th ed., pp. 806–853). Hoboken, NJ: Wiley.

Leinbach, M. D., Hort, B. E., & Fagot, B. I. (1997). Bears are for boys: Metaphorical associations in young children's gender stereotypes. *Cognitive Development*, *12*, 107–130. doi:10.1016/S0885-2014(97)90032-0

Liben, L. S., & Bigler, R. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development*, *67* (Serial No.269).

Liben, L. S., Bigler, R. S., & Krogh, H. R. (2001). Pink and blue collar jobs: Children's judgments of job status and job aspirations in relation to sex of worker. *Journal of Experimental Child Psychology*, *79*, 346–363. doi:10.1006/jecp.2000.2611

Liben, L. S., Bigler, R. S., Shechner, T., & Arthur, A. (2006, April). *Preschoolers' sex typing of self and others: Toward coordinated lifespan measures*. Poster presented at the Gender Development Research Conference, San Francisco, CA.

Liben, L. S., & Coyle, E. F. (2014). Developmental interventions to address the STEM gender gap: Exploring intended and unintended consequences. In L. S. Liben & R. S. Bigler (Eds.), *The role of gender in educational contexts and outcomes* (Vol. 47, pp. 77–115). San Diego, CA: Elsevier.

Liben, L. S., & Hilliard, L. J. (2010, October). *Preschoolers' gender vigilance: Effects of classroom organization*. Poster presented at the Gender Development Research Conference, San Francisco, CA.

Liben, L. S., & Signorella, M. L. (1980). Gender-related schemata and constructive memory in children. *Child Development*, *51*, 11–18. doi:10.2307/1129584

LoBue, V., & DeLoache, J. S. (2011). Pretty in pink: The early development of gender-stereotyped colour preferences. *British Journal of Developmental Psychology*, *29*, 656–667. doi:10.1111/j.2044-835X.2011.02027.x

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Martin, C. L., Eisenbud, L., & Rose, H. (1995). Children's gender-based reasoning about toys. *Child Development*, *66*, 1453–1471. doi:10.1111/j.1467-8624.1995.tb00945.x

Martin, C. L., Fabes, R. A., Hanish, L. D., & Hollenstein, T. (2005). Social dynamics in the preschool. *Developmental Review*, *25*, 299–327. doi:10.1016/j.dr.2005.10.001

Martin, C. L., Fabes, R. A., Hanish, L., Leonard, S., & Dinella, L. M. (2011). Experienced and expected similarity to same-gender peers: Moving toward a comprehensive model of gender segregation. *Sex Roles*, *65*, 421–434. doi:10.1007/s11199-011-0029-y

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development*, *42*, 1119–1134. doi:10.2307/1129498

Martin, C. L., Kornienko, O., Schaefer, D. R., Hanish, L. D., Fabes, R. A., & Goble, P. (2012). The role of sex of peers and gender-typed activities in young children's peer affiliative networks: A longitudinal analysis of selection and influence. *Child Development*, *84*, 921–937. doi:10.1111/cdev.12032

National Science Foundation. (2013). *Women, minorities, and persons with disabilities in science and engineering* (Special Report). Retrieved from http://www.nsf.gov/statistics/wmpd/2013/start.cfm?CFID=14614984&CFTOKEN=56461309&jsessionid=f0306d628c00285dc6a71c511a3461f5b7a7

National Science Foundation. (2014). *National Science Foundation Strategic Plan for 2014–2018*. Retrieved from http://www.nsf.gov/pubs/2014/nsf14043/nsf14043.pdf

Patterson, M. M. (2012). Self-perceived gender typicality, gender-typed attributes, and gender stereotype endorsement in elementary-school-aged children. *Sex Roles*, *67*, 422–434. doi:10.1007/s11199-012-0184-9

Petersen, J. S., & Hyde, J. S. (2014). Gender-related academic and occupational interests and goals. In L. S. Liben & R. S. Bigler (Eds.), *The role of gender in educational contexts and outcomes* (pp. 43–76). San Diego, CA: Elsevier.

Serbin, L. A., Poulin-Dubois, D., Colburne, K. A., Sen, M. G., & Eichstedt, J. A. (2001). Gender stereotyping in infancy: Visual preferences for and knowledge of gender-stereotyped toys in the second year. *International Journal of Behavioral Development*, *25*, 7–15. doi:10.1080/01650250042000078

Serbin, L. A., & Sprafkin, C. (1986). The salience of gender and the process of sex typing in three- to seven-year-old children. *Child Development*, *57*, 1188–1199. doi:10.2307/1130442

Shechner, T., & Lobel, T. E. (2010, April). *Gender development of children growing up in non-traditional families in Israel*. Poster presented at the Gender Development Research Conference, San Francisco, CA.

Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: A meta-analytic review. *Developmental Review*, *13*, 147–183. doi:10.1006/drev.1993.1007

Signorella, M. L., & Liben, L. S. (1984). Recall and reconstruction of gender-related pictures: Effects of attitude, task difficulty, and age. *Child Development*, *55*, 393–405.

Steinberg, S. (2004). The bitch who has everything. In S. Steinberg & J. Kincheloe (Eds.), *Kinderculture: The corporate construction of childhood* (pp. 207–218). Boulder, CO: Westview Press.

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development*, *85*, 1062–1076. doi:10.1111/cdev.12165

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*, 817–835. doi:10.1037/a0016127

Weinraub, M., Clemens, L. P., Sockloff, A., Ethridge, T., Gracely, E., & Myers, B. (1984). The development of sex role stereotypes in the third year: Relationships to gender labeling, gender identity, sex-types toy preference, and family characteristics. *Child Development*, *55*, 1493–1503. doi:10.2307/1130019

Weisgram, E. S., & Bigler, R. S. (2006). Girls and science careers: The role of altruistic values and attitudes about scientific tasks. *Journal of Applied Developmental Psychology*, *27*, 326–348. doi:10.1016/j.appdev.2006.04.004

Weisgram, E. S., Fulcher, M., & Dinella, L. M. (2014). Pink gives girls permission: Exploring the roles of explicit gender labels and gender-typed colors on preschool children's toy preferences. *Journal of Applied Developmental Psychology*, *35*, 401–409. doi:10.1016/j.appdev.2014.06.004

YoYoGames. (2012). *GameMaker for Mac*. Retrieved from http://www.yoyogames.com