We've Come a Long Way, Baby (But We're Not There Yet): Gender Past, Present, and Future

Lynn S. Liben

The Pennsylvania State University

Gender has long been, and continues to be, a powerful predictor of developmental experiences and outcomes. Observations drawn from personal history, developmental science, and life beyond the academy show that historically, gender constraints have diminished in some ways, but remain robust in others. Reviewed are children's constructive processes that—in interaction with the embedding ecology—foster the emergence and persistence of gendered phenomena. Reviews of interventions designed to increase girls' science participation demonstrate the need to evaluate both intended and unintended program consequences. Discussion of the single-sex schooling debate shows the importance of foundational conceptualizations of gender, and illuminates research-to-policy processes. After identifying newly emerging gender conceptualizations, the concluding section highlights the need to consider how gender conceptualizations do and should affect science and society.

The honor of being elected President of the Society for Research in Child Development (SRCD) carries with it not only countless organizational duties and the weighty responsibility of presenting an address at the biennial meeting, but also the reward of a sure-bet publication in *Child Development.* Liberated from the specter of editors and reviewers who normally haunt me as I write, I decided to sidestep the traditionally formal third-person genre of this journal and instead retain the character of my presidential address. I thus build arguments by drawing not only from mainstream conceptual and empirical work in developmental science, but also from personal history.

Importantly, as I learned as a Cornell undergraduate taking my first child development course from Urie Bronfenbrenner, personal experiences in the immediate social and physical environment (the microsystem) are spawned in dynamic interaction with contextually embedding layers (Bronfenbrenner, 1977). In an ecological systems model, the most distal and thus most encompassing layer is time, or what Elder (1998) referred to as the chronosystem. Thus, one reason to consider historical context is to enhance understanding of individual development.

But there is a second reason to highlight history, one that many of us learned in junior high school via the aphorism, "Those who cannot remember the past are condemned to repeat it" (Santayana, 1905, p. 284). Although revisiting this sentence in its original context shows that Santayana was actually critiquing "savages" for their failure to retain experiences rather than disparaging contemporaries for failing to heed history, it is the latter interpretation that I mean to suggest here. Examining historically distant contexts may reveal values, rules, events, or behaviors that seem surprising or even appalling when viewed through contemporary lenses. Such revelations provoke us—as developmental scientists—to ask whether similar phenomena still exist, and if so, to identify processes by which those undesirable social contexts are generated and maintained. Resulting insights may in turn suggest strategies for ameliorating unwanted effects and for avoiding their emergence in the first place. In short, the first reason for examining historical contexts is to further our scientific

understanding of individual development; the second is to suggest avenues for modifying social contexts in order to optimize developmental outcomes. Both goals resonate with SRCD's core mission—to advance developmental science and promote its use to improve human lives (SRCD, 2015).

My substantive focus is on gender development. In what ways are developmental processes and outcomes differentiated by gender? What accounts for the existence, importance, and persistence of gender-differentiated phenomena? What are their consequences for individuals and for society? How might answers to these questions inform developmental scientists' recommendations to parents, educators, and policy makers?

One way to reflect on how gender-related phenomena have evolved is by examining personal histories, and it is this tack that I take in the next section, *Gender Then and Now*. There I introduce my first thesis—encapsulated in the paper's title—that society has reduced, but not eliminated, gender-based constraints on human development. In the section entitled *Gender-Differentiating Processes*, I draw from developmental theory and research to identify processes that contribute to the persistence of gender constraints. In the next section, *Interventions*, I examine gender-related interventions designed to test hypothesized developmental processes and to reduce undesirable gender differences. In *Single-Sex Schools* I discuss the contemporary debate about single-sex education, a focus that allows a detailed look at one large-scale intervention strategy and an opportunity to consider science-to-policy issues more generally. In the final section, *Looking Back and Looking Ahead*, I recap the value of examining gender development across historical as well as ontogenetic time.

Before turning to personal history, illustrative research, sample interventions, and science-to-policy issues, I acknowledge explicitly that I draw my personal and research examples almost entirely from the United States, and within a narrow slice of the U.S. experience at that. Thus, although I expect that the arguments I make here are broadly applicable, this is indeed an expectation rather than an assertion. The current analysis (like most if not all others in our scholarly literature) has been developed within a particular context, and thus its applicability to diverse ecological systems must be tested rather than assumed.

I acknowledge, too, that my decision to retain the spirit of the spoken version of my presidential address in these pages has proven more challenging to implement than anticipated. In part, the challenge stems from the need to cull out all but a tiny number of my original 121 slides—a process that has convinced me more than ever that a picture is indeed worth a thousand words. I would therefore urge readers to click on the cited sources, particularly those leading to video clips and news articles. At least as challenging are differences in the tightness and detail of arguments that are prototypical of spoken versus written genres. In the former, it is routine to present a few-word newspaper headline or an isolated bar graph from a complex study; in the latter, it is normative to document contextual details, research design, and statistical findings. I have retained the style of the spoken genre to allow me to sample broadly from diverse kinds of evidence. I have also selected personal examples that necessarily emphasize experiences of women, and research examples that disproportionately represent my own program of work. My goal in offering diverse and self-relevant kernels is not so much to build, brick by brick, a tightly constructed edifice of a complete and universally generalizable argument. Rather, it is to demonstrate the value of approaching development by drawing not only on systematic, theory-driven empirical research, but also on personal experiences and on observations of the broader society that lies beyond the academy.

Gender Then and Now

An Earlier Generation

I begin my tale of personal history by recounting some of my mother's early career experiences. To set the stage (and to honor the memories of the two most influential members of my childhood microsystem), Figure 1 is a photograph of my parents, Florence Gettenberg Liben and Jay Liben, just after they were wed at New York's City Hall. Among the many qualities that made my mother special is that early in life she decided to become a physician, a goal that was neither common nor easy for girls of her time. After her high school graduation, my mother enrolled at Cornell University and followed the standard premedical curriculum. Although I have not unearthed memorabilia that shed light on how family, friends, or faculty responded to her chosen path, her transcript shows that she persisted, and early in her senior year she applied to medical school. One index of the difficulty women faced was a letter she received from the University of Pennsylvania's School of Medicine reproduced in Figure 2. The Dean's office thoughtfully wrote to suggest that she forgo a trip

Figure 1. Florence and Jay Liben, May 29, 1937.

from Ithaca to Philadelphia for an interview because they had acted unusually early and had already filled their year's quota of women.

Fortunately, my mother had also applied elsewhere, and in the fall of 1935, she matriculated in the New York University College of Medicine. A yearbook photograph of most of her class, reproduced in Figure 3, gives a rough idea of the gender distribution. A more precise count comes from her 1939 commencement program: Of the 124 names, 12 were women's. As an interesting aside, the list included Stella Chess, one of the originators of the concept of temperament, and Jonas Salk, the immunologist who developed the first vaccine against polio and who was—albeit less famously—a childhood friend of my parents.

Having completed medical school, women next encountered barriers in obtaining internships. Among my mother's saved papers are five pages of hand-written, 11-column spread sheets summarizing potential hospital internships along dimensions such as types of rotations (e.g., surgery, medicine, ob-gyn), numbers of internship slots, and hospital size. Also included were miscellaneous gender-related warnings such as a double-underlined comment about Beth David Hospital that read, "Rarely acc. [*accepts*] ♀."

Her applications led to an offer for a 2-year internship at New York Metropolitan Hospital. The inside address of her letter of appointment included her first name—Florence—but the salutation read "Dear Sir." Perhaps the greeting was only a typographical error. But perhaps it was indicative of the hospital's difficulty in accommodating women, an interpretation also suggested by the lack of sleeping accommodations provided for women interns on 24-hr ob-gyn shifts. There were two cots, but both were designated for men. Judging that an official request for facilities for women would likely be unsuccessful or unbearably slow, my parents took matters into their own hands. My mother donned her medical whites and borrowed a set for my (nonphysician) father. Sporting their official-looking garb and demeanor, they managed to enter the hospital with a folding cot, roll it past guards, nurses, and other staff, and install it in a small supply room that remained a facility for women interns until well after my mother's ob-gyn rotation had ended.

A Somewhat Later Era

My mother's records provide some sense of the 1930s and 1940s. My own memorabilia allow me to shed light on the decades that followed. Among
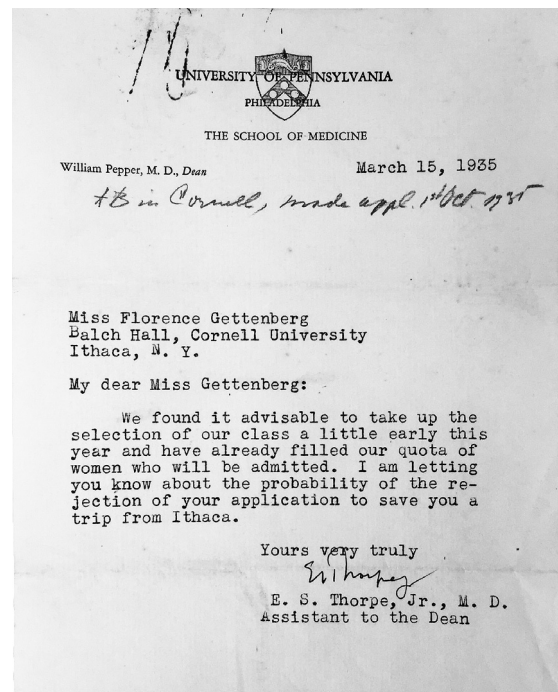

Figure 2. Letter from the University of Pennsylvania's School of Medicine to Florence Gettenberg regarding her application.

Figure 3. Class of 1939 of the New York University College of Medicine, as sophomores. Reproduced from the 1937 yearbook, *The Medical Violet*.

my stored treasures is a 45-page, purple-construction paper, ribbon-bound report I submitted in 1961 to meet a ninth-grade "Career Booklet" assignment. The exercise was intended to teach us about the impact of jobs on society, expose us to the importance of vocational planning, and give us each an opportunity to explore a career of potential interest. I selected law, probably seduced to it from watching weekly episodes of Perry Mason on our small, black-and-white television.

Among our assigned tasks was to interview someone in our selected field, and perhaps influenced by my mother's nontraditional career history, I interviewed a woman attorney. She did not prove to be an ideal feminist mentor. She warned me explicitly that law was no career for a woman, and recommended against my pursuing it. The typed report I submitted included similarly discouraging generalizations, beginning on page 2 when I commented that, "It is well known that men definately [*sic*] hold almost all legal jobs" and then reported that estimates of the proportion of women lawyers ranged between 2% and 3.5%. Some feminist optimism did emerge in my writing, however, when I cheerfully noted that, "Despite handicaps, women have, through hard, tedious, work, risen in their field" and continued with almost palpable pleasure that, "Such jobs as assistant attorney-general, judges of state supreme courts, and other lower judges have all been filled by women."

Only a few years later, in 1964, Title VII of the Civil Rights Act became law, thereby prohibiting discrimination in employment and job training on the basis of race, color, religion, sex, and national origin. But even as this legal act marked societal progress, more subtle and perhaps therefore more pernicious gender-based messages about jobs and

life's roles were communicated. The year that Title VII was passed, I was a senior in high school and, because I was a girl with an English grade of at least 85, I was allowed to (and, in fact, was required to) take the national test used to identify the Betty Crocker Homemaker of Tomorrow. (Even then the irony was not lost on me; I remember noticing that virtually none of the girls enrolled in any of my high school's domestic arts classes met the English grade requirement.)

I probably would not have included this particular example were it not for the fact that I received the highest test score in my school. My prize was a small heart-shaped gold pin on which was embossed a burning hearth. But more thrilling was the congratulatory letter I received from Betty Crocker that is reproduced in Figure 4. The closing sentence reads, "The qualities you have shown are the best possible foundation for a happy and successful future as a real homemaker--the most important career a woman can have." In short, among the messages that I and others of my cohort received in our time was that our education, careers, and domestic roles were meant to follow gender-specialized paths. I was *not* supposed to become a lawyer, but I *was* supposed to become a "real" homemaker.
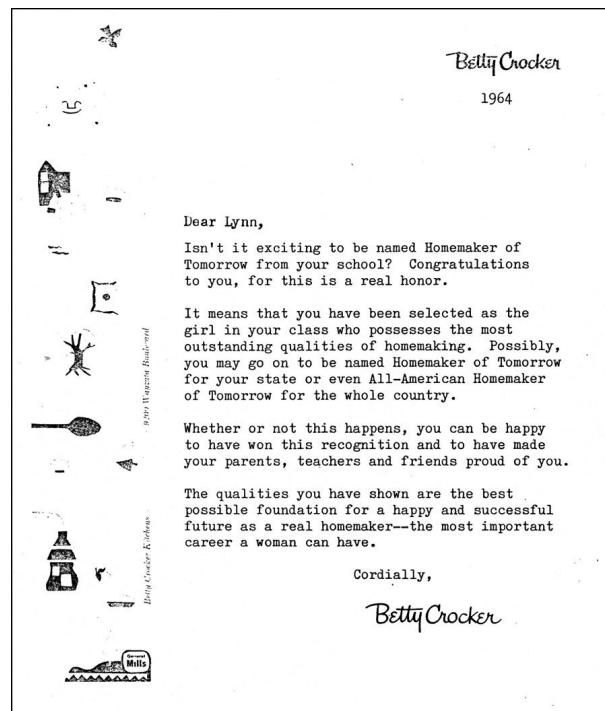

Figure 4. Congratulatory letter from Betty Crocker to Lynn Liben, 1964.

It was also during the 1960s that the phrase I used for the title of the current article became famous. It emerged as part of a pervasive print and television advertising campaign for a new brand of cigarettes called *Virginia Slims.* Actually, the inspirational phrase was not, "*We've* come a long way, Baby" but rather, "*You've* come a long way, Baby." The "you" referred to women and the "long way" to progress on women's rights. A collection of advertisements may be found online (see Virginia Slims Commercials, 1969). One, for example, set in an earlier era, begins as a man is returning home from work. He is greeted at the door by an aproned woman who takes his coat and hat and kisses his forehead. The man (presumably her husband) then settles into his easy chair. During these scenes, the male voiceover says, "Back in the old days, men were the masters, women were the slaves. Women had the duties, men had all the rights." The next scene shows suffragettes marching for the vote, accompanied by the voice saying, "Then, at last, women won their rights. And then, one by one, they won them all." The next series of scenes shows a contemporary woman, who—accompanied by the song *You've Come a Long Way, Baby*—applies eye makeup and lipstick, pins up her hair, and slings a boa scarf over her shoulder. The final series of images shows the woman seductively selecting, lighting, and smoking a Virginia Slim, and then walking off into the distance. These final scenes are accompanied by a reprise of the song, intermingled with a woman's voice noting that Virginia Slims are "slimmer than the fat cigarettes men smoke" and recounting other qualities that make them particularly appealing to women.

Contrasting Then and Now

My decision to share some personal experiences from my mother's and my lives was motivated by the belief that it is informative to see how overarching societal constraints play out in individual lives. My hope was that those who lived through the same era would be reminded of similar experiences and those who came of age more recently would get a better sense of the lived experiences of those earlier times. The anecdotes I chose were intended to depict circumstances sufficiently unlike today's that they would lead to the conclusion that society has, indeed, come a long way.

Many systematic empirical data are consistent with such a conclusion. Returning to the domain of medicine, for example, the U.S. medical school enrollment during 2014–2015 was reported by the Association of American Medical Colleges (2015) to be 47% women, a far cry from the 10% of my mother's class. However, detailed data on specialties (Association of American Medical Colleges, 2014) reveal continuing signs of gender imbalance. For example, women now make up less than 10% of the high-status specialties of vascular surgery, vascular and interventional radiology, neurological surgery, interventional cardiology, urology, thoracic surgery, and orthopedic surgery. The only specialties in which women outnumber men are pediatrics (60.4%), ob-gyn (51.8%), pediatric internal medicine (50.9%), and child and adolescent psychiatry (50.4%). There are also striking gender differences in financial compensation. For example, 2012 Medicare payments averaged $118,782 for men and $63,346 for women, differences that can be accounted for only in part by gender-differentiated patterns of specialization (Pratini, 2014).

Similar patterns of change and persistence in gendered outcomes are found in the more general domain of STEM (science, technology, engineering, and math). Illustratively, between 1966 and 2006 the proportion of bachelor's degrees awarded to women rose to or beyond parity in some fields (e.g., in biological and agricultural sciences, from 25% to 60%; in chemistry, from 19% to 52%; in Earth science, from 9% to 41%), but in others, despite increases, women's participation rates remain far below men's (e.g., in computer science, the change was from 15% to 21%; in physics, from 5% to 21%; and in engineering, from 1% to 19%). Similar gender differences are still seen in the pre-college years as measured by participation and success on Advanced Placement tests, in STEM-related academic competitions, and on other national and international educational assessments (see Leaper, 2015; Liben, 2015a; Liben & Coyle, 2014).

The pattern of reduced, but nevertheless persistent gender differences appears in domains as diverse as athletics, politics, parenting, and business, to name only a few (e.g., see reviews of gender development by Blakemore, Berenbaum, & Liben, 2009; Hines, 2015; Leaper, 2015; Ruble, Martin, & Berenbaum, 2006). In part—as argued in detail later—these contemporary gender-differentiated outcomes are supported by individuals who explicitly hold and promulgate deep-seated beliefs about inherent, pervasive, and desirable gender distinctions. But they are also sustained by more general ontogenetic processes of individual development. In the following section, I discuss theoretical and empirical work that addresses

developmental processes entailed in sustaining these long-standing gender distinctions.

Gender-Differentiating Processes

Gender Schemata and Information Processing

My own foray into the study of gender development was motivated largely by an interest in extending my dissertation from the cognitive to the social realm. Like many developmental dissertations of the time, my work was designed to test hypotheses derived from Piagetian theory. My explicit goal was to test Piaget's proposal that memory is a constructive process such that what a child perceives and remembers of a stimulus or event is determined largely by the child's underlying conceptual schemes. In my initial research (Liben, 1974), I studied elementary school children's immediate and delayed reproductions of a drawing of a half-filled, tipped bottle. As predicted, data showed significant (albeit imperfect) links between children's assessed spatial concepts and their success in reproducing the horizontal water line of the original stimulus. Although I continued to pursue related spatial questions (and still do so; see Liben, 2014), I also became interested in seeing if similar processes would operate in social domains. To that end I embarked on a series of studies in collaboration with Margaret Signorella who was then a student in my graduate seminar on memory development. In our work (Liben & Signorella, 1980, 1993; Signorella & Liben, 1984, 1985) we asked children to remember drawings that depicted people engaged in either traditional jobs or activities (e.g., a male construction worker; a female nurse) or nontraditional ones (e.g., a female dentist; a male librarian). Studies varied along dimensions such as the type of memory task (recognition vs. free recall), task difficulty (e.g., varying speed of item presentation, number of items, or delay intervals), and whether verbal labels accompanied each picture.

Consistent with the hypothesis of schema-influenced memory, children remembered significantly more traditional than nontraditional pictures. Furthermore, although children's recall of traditional items was nearly always true to the original stimulus, recollections of nontraditional stimuli were distorted in ways that made them traditional (e.g., recalling a male secretary as a typewriter repairman or a woman judge as a school cafeteria worker). Data were also consistent with the prediction that what determines memories are individual children's gender schemata rather than mere exposure to actu-arial realities in the surrounding culture. That is, the finding that memory was significantly better for traditional than nontraditional items was stronger (and sometimes found only) among children who endorsed cultural gender stereotypes on a separate attitudinal measure.

Evidence of constructive memory processes like these can help to account for the perpetuation of gender-differentiated beliefs and behaviors over the decades. Consider, for example, a young girl who visits a dental office and encounters a female dentist. If the child forgets the encounter entirely or reinterprets the dentist as a dental hygienist, she is unlikely to revise her already-established gender schema that includes the belief that dentists are necessarily men. Her belief that dentistry is not for people like her would affect her beliefs about her later career options, and might have immediate effects on the degree to which she views related skill-building play and school lessons as self-relevant (e.g., Eccles, 2014).

Gender Schemata and Self-Selected Experience

Gender schemata are powerful not only in influencing how children interpret and recall new experiences, but they are also powerful in influencing what experiences children have in the first place. A seminal model addressed to the role of gender in patterns of children's self-driven engagement is gender schema theory (GST; Martin & Halverson, 1981). Central to this model is the proposal that children use gender schemata to guide their interaction with the surrounding environment. A key component of those schemata is information about how the culture stereotypes objects and activities. Confronted with some object, the child categorizes it as "for girls" or "for boys," and then, depending on the child's own gender identity ("I am a girl" or "I am a boy"), either approaches or avoids it. So, for example, a girl who sees a box of interlocking blocks might identify it as a toy stereotyped "for boys," and, because she identifies as a girl, avoid playing with it. Empirical tests of this theoretical formulation have provided support for the importance of the child's belief about whether something is meant for boys or girls. For example, children say they would prefer to play with completely novel toys that have been labeled as for their own gender, even when the toys that had been labeled for the other gender are more attractive overall (Martin, Eisenbud, & Rose, 1995).

In and of itself, engagement with a particular toy or activity may not appear to be terribly important,

but compounded over opportunities and across years, such choices influence the emergence and expansion of interests and skills that can matter a great deal in producing and sustaining the gender-differentiated developmental outcomes discussed earlier. For example, if girls systematically avoid toys such as Lincoln Logs, Tinker Toys, or model airplane sets, they lose opportunities to build spatial thinking and mechanical reasoning skills recognized as foundational for STEM fields, including some (like engineering) that remain male dominated (e.g., Liben & Coyle, 2014).

In a later constructivist approach to gender development—the dual pathways model (DPM)—Rebecca Bigler and I likewise argued for the importance of gender schemata in guiding children's gendered behaviors and development (Liben & Bigler, 2002). DPM, however, assigned a more explicit role to individual differences. That is, although we concurred with Martin and Halverson's (1981) observation that virtually all young children have acquired knowledge about what the culture defines as "for girls" and "for boys," we proposed that also important was the impact of individual differences including those in (a) the degree to which children endorse (not merely know) cultural gender stereotypes (the *gender schema filter*), (b) the extent to which children routinely process experiences through the lenses of gender (the *gender salience filter*), and (c) interests and talents that exist aside from gender-related constraints (the *interest filter*). In addition, we posited the operation of two pathways. The first—the *attitudinal pathway*—predicts that gender-stereotyped attitudes about others drive patterns of engagement for the self (i.e., an other-to-self pathway). The second—the *personal pathway*—posits the simultaneous contribution of a reverse process in which children's own interests and behaviors influence their attitudes about what is gender appropriate for others (i.e., a self-to-other pathway).

Empirical findings have demonstrated the value of studying effects of individual differences and dual pathways. For example, as discussed in greater detail later, findings from an experimental study (Coyle & Liben, in press) showed that under some conditions, preschool girls respond differently to a game about occupations in relation to individual differences in their general attentiveness to gender. Findings from a longitudinal study (Liben & Bigler, 2002) showed support for the hypothesized impact of the personal (self-to-other) pathway: Middle school boys who initially (fall of Grade 6) selected more feminine descriptors as self-descrip-

tive later (spring of Grade 7) reported more egalitarian gender attitudes about others.

How Gender Comes to Matter

Thus far I have discussed work demonstrating that gender schemata have a powerful influence on the way that new information is processed and remembered, on children's patterns of experiential engagement, and on the ways that gender schemata are applied and modified over time. Given the power of children's cognitions about gender, it is important to step back and ask how these cognitive gender schemas emerge in the first place, and more generally, why gender is such a salient and important dimension along which things such as people, toys, activities, roles, and jobs are categorized. In short, how do gender-based categorizations come to take on such importance and have such far-reaching effects?

The answer I offer here stems from collaborative work with Rebecca Bigler in which we formulated developmental intergroup theory (DIT; Bigler & Liben, 2006, 2007). Our theory is rooted in research on constructive processes in gender development such as the work just described, and in Bigler's research on intergroup processes (e.g., Bigler, 1995; Bigler, Brown, & Markell, 2001; Bigler, Jones, & Lobliner, 1997). Our formulation begins with the observation that there are many attributes that differ across people, and with the position that humans as a species are not hardwired to attend to any one or two of these attributes in particular (e.g., biological sex or skin color). We have proposed, instead, that humans actively apply general and flexible cognitive systems to infer what human attributes count as important within their ecological niche. Furthermore, we argue that certain factors affect the likelihood that a particular attribute will become psychologically salient, including whether the candidate attribute is visibly distinctive (*perceptual discriminability*), whether different language is used for people who differ on that attribute (*explicit labeling*), whether the attribute is explicitly used to assign or permit some behavior or opportunity (*explicit use*), and whether the attribute systematically covaries with some other observable condition or attribute even in the absence of an explicit rule or explanation of why it does so (*implicit use*). Also enhancing the psychological salience of an attribute are strikingly uneven (rather than balanced) distributions of subgroups defined by a given attribute (*proportional group size*).

As explained in more detail in the original presentation of the theory (Bigler & Liben, 2006), we argue that once attributes have become psychologically salient, they are likely to be used to form social group categories because of the general tendency to reduce cognitive complexity through categorization (see Allport, 1954). Having these categories then leads children to search for (or invent) links between the category and other qualities (i.e., social group stereotypes). Having categories also leads children to identify with one group (in-group) rather than the other (out-group), and self-affirming motivations lead to valuing the former and disparaging the latter (in-group favoritism and out-group bias).

When considered in light of the factors of perceptual discriminability, explicit labeling, explicit use, and implicit use, it is easy to see why gender is a likely basis for social-group categorization and stereotyping. Boys and men look different from girls and women; much language is gender specific (e.g., waiter, waitress; he, she; Nick, Nicole); gender is often explicitly used to sort people (e.g., Boy Scouts, Girl Scouts; boys' soccer, girls' soccer); and there are many cases in which gender is observably correlated, without explanation, to something else (e.g., all U.S. presidents to date have been men). Importantly, in our discussions of DIT and its implications for social policy (e.g., Bigler, 2013; Bigler & Liben, 2006; Liben & Bigler, 2015; Liben, Bigler, & Hilliard, 2014), we have argued that societies may either minimize or exaggerate these differentiating factors.

As an experiential demonstration of the power of societal variations in amplifying versus minimizing societal gender distinctions, consider the relative ease or difficulty of identifying the women soldiers in the two photographs shown in Figure 5. In the photograph on the left, the task is easy in view of gender-distinct clothing and hair styles. In the photograph on the right, the task is far more difficult given the traditionally masculine dress and hairstyles of everyone in the image. And indeed it is a testament to the power of such paraphernalia that most viewers are taken aback to learn that every one of the soldiers pictured in the photograph on the right is a woman.

As an empirical demonstration of the power of social conditions in generating gender stereotypes and prejudices, consider the findings from a study designed explicitly as a test of DIT (Hilliard & Liben, 2010). We tested whether increasing teachers' explicit labeling and use of gender in preschool classrooms would—as the theory predicts—lead to increases in children's gender-stereotyped attitudes and to increases in group bias (i.e., favoring one's own gender group or avoiding the other). Using a classroom procedure developed by Bigler (1995), we asked preschool teachers assigned to the experimental condition to use gender-specific rather than gender-inclusive language (e.g., "Good morning, boys and girls" rather than "Good morning, children") and to use gender to structure classroom activities (e.g., "Could I have a girl to help pass out the paper?" or "OK, boys, please line up for lunch . . . now girls, please line up"). Teachers were, however, explicitly cautioned against using gender in a comparative or competitive manner (e.g., to avoid requests such as "Let's see who can clean up faster, the boys or the girls"). Teachers assigned to the control condition were asked to maintain the established gender-neutral language and structure of the school. Periodic observations showed that teachers' behaviors indeed differed between the two conditions.

Figure 5. Soldiers in military dress. Left: Uniforms of the United States Army, 2008, from the United States Army Institute of Heraldry, from Wikimedia Commons. Right: Cadets at Sandhurst. UK Photo by Tim Graham/Reproduced courtesy of Getty Images.

Both prior to and following 2 weeks of this classroom manipulation, children were given measures of their gender-stereotyped attitudes and of their interest in playing with each of the children in their class. In addition, they were observed during regularly scheduled free-play periods to record with whom they played. Findings revealed an identical pattern of results on all three dependent measures: Children in the control condition showed no significant change between pre- and posttest, whereas children in the experimental condition did. Specifically, children in the experimental condition showed a significant increase in their endorsement of cultural gender stereotypes, a significant decrease in their reported interest in playing with classmates of the other gender, and a significant decrease in actual play with children of the other gender during free-play periods. An illustration of the data pattern is provided in Figure 6. As discussed earlier, an environment that strengthens endorsement of gender stereotypes reduces children's engagement in other gender activities, in turn reducing children's opportunities to develop a wide set of skills and interests. The finding that the experimental classrooms also led children to avoid play with children of the other gender likewise has long-term implications given that there is a compelling literature showing that children who play more with children of their own gender become increasingly more gender differentiated in both social and academic domains (e.g., see Martin & Fabes, 2001; Martin, Fabes, & Hanish, 2014).

Gender-Relevant Interventions

My goal in the prior section was to highlight processes that contribute to establishing and maintaining the content and consequences of gender schemata and stereotypes even in the face of countervailing forces such as antidiscrimination laws.
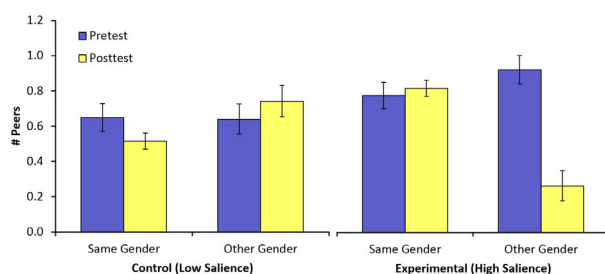

Figure 6. Number of peers played with, on average, during 15-s observational units by condition, peer type, and time. Based on Hilliard and Liben (2010), with permission of the authors.

My goal in the current section is to focus on interventions that may diminish the power of gender schemata. I begin by discussing researcher-initiated interventions. These are designed primarily as tests of processes identified in theoretical work, but the research also provides proof-of-concept demonstrations that may later be useful for designing large-scale intervention programs. I then discuss field-generated interventions and innovations that are motivated primarily by the practical goal of reducing existing but unwanted gender distinctions in behaviors and aspirations.

Researcher-Initiated Interventions

The three researcher-initiated interventions I describe here again represent collaborative work with Rebecca Bigler. The first (Bigler & Liben, 1990) was designed to test the effect of an intervention aimed at changing the content of children's gender schemas. Our reasoning was that if children's schemas about what men and women can do could be broadened, children would be better prepared to remember counterstereotyped material and would likewise be more likely to view a broader range of activities and roles as self-relevant.

Elementary school children were first given a measure that assessed the strength of their gender stereotypes, and based on their scores, divided into two matched groups. Children in both control and experimental groups received five 20-min classroom lessons about interests and training needed for various jobs. Children in the experimental group were also taught explicitly that gender was irrelevant for determining who could do a job. Because we anticipated that constructive processes would make it difficult for children to understand and remember lessons that included counterstereotypic information, we designed the intervention so that we could monitor children's understanding, and correct it when necessary. Thus, we worked with children in small groups and posed questions to elicit oral responses, a format that provided opportunities for immediate instructor feedback. For example, after hearing rules about the necessity of job skills and interests but the irrelevance of gender, children were told that Ann likes to build and knows how to drive a bulldozer, and then asked, "Could Ann be a construction worker?" and "How do you know?" Some children responded incorrectly in ways that suggested distorting effects of children's gender stereotypes. For example, one child answered "No, because Ann is a girl" and another answered "Yes, because he [*sic*] followed the rules."

In cases like these, there was an additional review of the material (including the rule that gender is not a relevant job criterion) and additional questions were posed.

Following all lessons, children attended story sessions on each of the next 12 school days. Each story involved a character who was engaged in a culturally stereotyped occupation (e.g., a dentist). Half were presented in a stereotypic version and half in a counterstereotypic version (in this example, a male or female dentist, respectively). In individual interviews that followed, children were asked questions about the story in a way that would elicit gendered pronouns and possessives to refer to the critical character. In that way, we could infer the child's memory of the critical character's gender. Findings showed that pronoun use was significantly more often flawed when questions concerned counterstereotypic characters. More importantly from the perspective of the intervention, children in the experimental group were significantly more accurate in recalling counterstereotypic stories than were children in the control group. There was, however, no indication that the experimental intervention had affected children's personal job interests.

In a second intervention (Bigler & Liben, 1992), we targeted a more general reasoning skill—double classification. We reasoned that children who have difficulty classifying something along two dimensions simultaneously would find it difficult to process counterstereotypic stimuli (e.g., a woman engineer) because such stimuli require that the child process the intersection of two categories simultaneously (ENGINEER + FEMALE). We reasoned that stereotypic stimuli (e.g., a male engineer) would not require the same advanced level of classification because such items would be processed as a single, integrated unit (because ENGINEER automatically entails MALE). To test this notion we gave some (but not other) elementary school children practice sorting drawings into 2 × 2 matrices, using rows to sort along one dimension and columns to sort along another. For example, given a scrambled pile of drawings of purple- and orange-colored shoes and hats, the child might sort footwear and headgear into different rows, dividing purple and orange items into different columns. As hypothesized, children who had learned to perform double classification tasks were significantly better at remembering nontraditional stories than were children who had not been so trained.

The third illustrative intervention (Lamb, Bigler, Liben, & Green, 2009) was designed to teach chil-

dren to recognize and confront peers' sexist messages. The goal was thus not only to modify the content or use of individual children's gender schemata (as in the prior two studies), but also to reduce the occurrence of sexist behaviors and thus modify the surrounding ecology. On each of 6 school days, young elementary school children were given 20-min classroom lessons about a type of sexism (e.g., exclusion of an out-group member) and were taught an appropriate and verbally catchy retort (e.g., "You can't say girls [boys] can't play!"). Based on the hypothesis that children's own behaviors will influence the attitudes they develop about others (see the earlier discussion of the personal pathway in DPM), children in the "practice" condition were taught response strategies by asking them to create and participate in skits illustrating each of the sexist remarks and the paired retort. Children in the "narrative" condition were taught about sexist remarks and retorts through instructor-prepared stories.

Both prior to and following the completion of all six lessons, children were given a series of vignettes in which someone displays a sexist behavior, and participant children were asked how they would respond to the incident. At the close of the study, children were also tested behaviorally. Specifically, each child was asked to take a gender "inappropriate" object to the office (for boys, it was a pink purse; for girls, a tool belt). Along the way, the child was stopped by a confederate peer who made a scripted sexist remark about what the child was carrying. The dependent measure was whether the participant child challenged the confederate, and if so, in what way.

The intervention was successful with respect to its targeted goal of teaching children to recognize and respond to peers' sexist behaviors, with significantly stronger effects evident in the practice condition. In responding to vignettes at pretest, children almost never reported that they would challenge peers' sexist remarks: Only 10% of the children stated that they would challenge even one sexist remark, and these children averaged only 4 challenges each (of a possible 13). By the posttest, the incidence of reported challenges was strikingly higher. For example, in response to vignettes that depicted a child making a disparaging comment about another child's counterstereotypic appearance or behavior (e.g., a vignette in which a girl was asked, "Why do you have a boy's haircut?"), 78% of children in the practice group and 33% of children in the narrative group reported that they would challenge the child who had asked the sexist

question. Although on the behavioral measure retorts were rare, almost all (10 of 11) were produced by children in the practice condition.

Furthermore, at a delayed posttest 6 months later, the greater frequency of challenges to sexist remarks on the vignette measure had spread not only to additional children in the narrative condition, but also to children who had not participated in the intervention at all (because they had enrolled in the school after the classroom intervention lessons had been completed). The latter finding is particularly encouraging insofar as it suggests that even a relatively brief intervention that addresses peers' sexist behaviors can lead to widespread and relatively long-lasting changes in the broader ecological system.

In summary, the interventions reviewed above provide support for a number of theoretical hypotheses about gender schema-based processing, show that even short interventions can expand children's attitudes about what people—irrespective of gender—can or should do, and can enhance children's success in remembering nontraditional material. Given the theoretical frameworks discussed earlier, these conditions would—over the long term —be expected to broaden children's own interests and activities, although the data showed little evidence that children's personal interests were immediately affected by the interventions.

Field-Initiated Interventions

I now turn to interventions that have been designed and implemented in the field rather than in research laboratories. I draw my examples from innovations or programs intended to increase girls' participation in traditionally masculine activities that have been hypothesized or demonstrated to affect children's attraction to, and success in, STEM (see also Liben, 2015a; Liben & Coyle, 2014). Interventions in, this arena are of particular interest because they hold promise not only for expanding the opportunities and life satisfactions of individuals, but also for expanding the size and diversity of the societal workforce.

The cartoon reproduced in Figure 7 identifies two STEM intervention approaches. The first is to repudiate the notion that the target activity (here, math) is unfeminine; the second is to repackage the target activity to render it feminine (here, making it pink). The cartoonist has labeled the second of these approaches as the common one, and indeed, it is easy to find many field-initiated innovations that instantiate this approach.

To illustrate, I return to the comment I made earlier in the course of describing GST, "if girls systematically avoid toys such as Lincoln Logs, Tinker Toys, or model airplane sets, they lose opportunities to build spatial thinking and mechanical reasoning skills recognized as foundational for STEM fields" (p. 11). As one way of attracting girls to play with (and, presumably, to buy) toys like these, toy manufacturers have, literally, made them pink. The LEGO group, for example, has produced sets of their iconic interlocking blocks (bricks) in colors explicitly intended to appeal to girls. Illustratively, an advertisement for the LEGO *Pink Brick Box* (Amazon, 2015) describes it as having "bricks in colors you love!" Comments from consumers that appear on the Amazon Web page are consistent with this marketing claim. For example, "Laura" (on March 13, 2012) writes, "My grand daughter loves this girlie set! She's 4 and prefers playing with this set now, rather than her brother's black, grey and red Lego sets, the ones she used to spend hours playing with." Similarly, "an active mom" (August 23, 2010) writes, "My daughter & son fought over his legos. she was thrilled to have her own. we're so happy they have girl Legos!!" (As an aside, it is interesting to note that apparently neither of these girls had earlier spurned unpink LEGOs.)

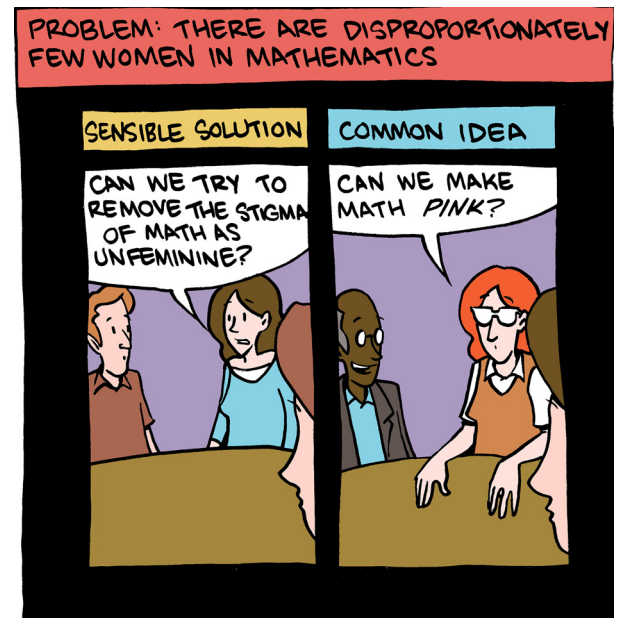It is not only color hues, though, that make some LEGO sets "pink" (i.e., feminized). Marketing for

Figure 7. Two approaches to the STEM gender gap. Reproduced with permission of Zach Weinersmith, http://www.smbc-comics.com/index.php?db=comic&id=1962#comic.

these sets also emphasizes traditionally feminine role play and story narratives, thus seemingly encouraging the use of LEGOs for symbolic play rather than for the construction activities thought to be conducive to developing STEM-relevant skills. For example, even the general-purpose Pink Brick Box set is advertised as offering a "special mix of elements and bricks with girl appeal" providing a "classic theme focus for role play." Specialty sets are even more directly targeted to stereotypically feminine interests and play styles. An advertisement for the LEGO *Butterfly Beauty Shop*, for example, urges girls to "Get primped and pretty" and to "Shop for lipstick, makeup and hair accessories! Emma and all of her friends will look fabulous with bows, sunglasses, a hairbrush, mirror, lipsticks and new hair styles" (LEGO Shop, 2015). It is especially compelling to view these descriptions in full to see the toys, the way that they are marketed, and the consumer comments that they elicit, and thus I urge readers to examine the cited (and related) Web sites.

In addition to reconfiguring classic toys as "pink," designers are also developing entirely new toys and marketing campaigns to attract girls and women to STEM. A recent example is the line of *GoldieBlox* toys designed by engineer Debra Sterling. In a TED Talk describing what led her to develop this line of toys, Sterling (2013) begins by describing her own checkered entry into engineering and sharing some challenges she experienced on the way to completing her Stanford engineering degree. She reports that years after her graduation, she became dismayed to realize that the toys associated with developing strong spatial skills "have been marketed to boys for over 100 years . . . [and] meanwhile all we got are the dolls and makeup kits." In various media, she explicitly describes the company she founded as "a toy company out to inspire the next generation of female engineers" (Sterling, 2015).

The toy series centers on Goldie, a female character who engages in solving various engineering problems. In the first set of the series—*GoldieBlox and the Spinning Machine*—Goldie assembles increasingly complex belt drives that enable her to spin, first, her dog, and then, multiple animals. The boxed toy set includes a book that narrates Goldie's motivation for building the spinning machine and provides instructions for assembling it. The set also contains all necessary pieces such as pegs, wheels, washers, and a ribbon (which serves as the belt) and the plastic animals.

There are multiple ways in which this set is "pink." Goldie initially becomes interested in the problem by finding a spinning ballerina in her mother's jewelry box, thereby introducing the engineering challenge via two classically feminine domains—ballet dancing and jewelry. Additionally, the toy set engages girls via reading, an activity often identified as feminine (although see J. Hyde, 2005; J. Hyde & Linn, 1988 for an alternative perspective). In her TED Talk, Sterling (2013) explicitly discusses her intentional use of reading to attract girls to *GoldieBlox*. She reports that when she initially gave some girls prototypes of her construction toys, they quickly became bored. When she asked them to tell her about their favorite toy, many said that they loved reading, ran off to get a book, and asked her to read the story with them. In what Sterling labels as "a really simple aha idea," she decided to "put those two things together: Spatial plus verbal, a construction set plus stories." Her next comment suggests her awareness of both approaches depicted in the cartoon of Figure 7, "And what if those stories were about a girl engineer character named Goldie Blox . . . it would bring in a role model and it would bring in the narrative that they so love." In other contexts, too, Sterling promotes the message that there is no incompatibility between being feminine and being an engineer. For example, on a Web site describing her personal history and company (Sterling, 2015), she features a fan letter that makes this point:

> . . . It really is inspiring to hear what you have done and are doing as I find people surprised that I can find Engineering interesting and still like to dress up or get my hair done. I'm glad that someone is trying [to] tell society that it doesn't come down to choosing between two things, but being able to have the best of both worlds.

Another program that is even more obviously designed to convey the message that science and femininity are not incompatible is *Science Cheerleader* (Cavalier, 2015). In this program, former National Football League and National Basketball Association cheerleaders who now work in STEM fields travel to various venues (e.g., sports events, fairs) to present cheerleading performances and encourage participation in science. In their pom-pom routines (best appreciated by viewing the video *Watch the Science Cheerleaders* found on the home page of Cavalier, 2015), these women cheer in favor of science and against gender stereotypes, chanting, for example, "We're bustin'—we're bustin'—we're bustin' down the stereotype! . . .

Gooooo SCIENCE!" The video shows cheerleading routines, brief interviews with cheerleaders about their science accomplishments, and footage of audience reactions. For example, in response to being asked "what do you want to be?" a young girl responds (pom-poms in hand), "a doctor, a teacher, and a cheerleader."

Another intervention that appears to be intended to combat the notion that science and femininity are incompatible is a film that was produced by the European Commission as part of their campaign, *Science: It's a Girl Thing!* This is another case in which viewing the 1-min video is far more powerful than any verbal description I could offer, but, in brief, the video shows a series of images of three high-heeled, highly attractive women strutting and posturing seductively to fast-moving, high-tech music. These scenes are interspersed with "science" scenes including several that involve the exchange of provocative glances between the women and a male scientist (initially shown peering into a microscope, but later ignoring it, his attention now riveted on the women); pictures of laboratory vessels (e.g., Petri dishes and beakers); images of cosmetics (including lipstick and dripping nail polish); a chemical ball-and-stick molecular model and then floating red, white, and blue balls; and a large letter H accompanied by the word "Hydrogen." In the closing scenes, the women don protective eyeglasses which they continue to wear as the title of the campaign appears, one word at a time. In the word "Science," the letter "I" is represented by an upright tube of lipstick.

Unfortunately from the perspective of archival scholarship, this film elicited sufficient outrage that is was pulled from the European Commission Web site, although the general campaign continues (European Commission, 2015). As of the writing of this article, the original film can still be viewed in a critical online article by Collins (2012), and a brief clip appears in an online news show in which the two news anchors discuss their differing reactions to the film (The Young Turks, 2012). The original film also inspired multiple parodies. One that is refreshingly funny to me (although perhaps unacceptably offensive to others) was produced by a group of Bristol University neuropsychology graduate students (Peacock, 2012).

Far more research is needed to evaluate the impact of programs like these. First, it would be essential to test whether the intended lessons are indeed learned. For example, does playing with Goldie or seeing a performance of the Science Cheerleaders successfully reduce children's gender stereotyping of STEM fields? Does it increase girls' interest in becoming an engineer or entering some other STEM occupation? Second, and as argued in greater detail elsewhere (see Liben, 2015a; Liben & Coyle, 2014), it is also crucial to test for unintended consequences, or what I have more provocatively referred to as collateral damage (Liben, 2010). One can well imagine that there would be unintended lessons about what it means to be a woman when the models who tout science are women who might be labeled in scholarly terminology as "hyperfeminized" (Coyle & Liben, in press) or in common parlance as "sexy." What will men who watch the *Science: It's a Girl Thing!* video think about the contributions women colleagues will make to collaborative projects? Might those thoughts influence their subsequent interactions with those women colleagues? Do such programs increase the psychological salience of gender in schools and in the workplace? Are girls who view presentations like the *Science: It's a Girl Thing!* video led to ruminate on how to enhance their analytic thinking skills or how to enhance their sexual attractiveness?

Although I know of no research that addresses these or similar questions about the field-initiated interventions just discussed, findings from two recent laboratory studies buttress the claim that such research is sorely needed. In one relevant study, we (Coyle & Liben, in press) invited preschool girls into the laboratory to play a computer game in which a female character engages in various jobs, including traditionally masculine ones such as chemist. We were interested in studying the impact of game play in relation to two variables. One was the character who enacted the jobs—either a hyperfeminized *Barbie* or a less feminized Playmobil character named *Jane.* The second was an individual-difference variable identified in the DPM —the degree to which the child routinely processes experiences through the lenses of gender (i.e., the gender salience filter or GSF).

Findings showed that game play led to no increase in girls' interests in masculine jobs, irrespective of game character or girls' GSF. However —and relevant to the suggestion that there may be unintended effects—playing the game did affect girls' interest in traditionally feminine activities in interaction with the two variables we studied. Specifically, when Jane was the game character, girls' interests in traditionally feminine activities did not change, irrespective of GSF. However, when Barbie was the character, those girls who routinely pay high levels of attention to gender became even more strongly interested in traditionally femi-

nine activities after playing the game. In short, "girly girls" who played the occupational game with Barbie—even a Barbie who enacted a broad range of jobs—became yet more girly as a result of the experience. It seems reasonable to expect that for some girls, exposure to *Science Cheerleaders* would work similarly, that is, having little impact on their STEM aspirations, but significant impact on their cheerleading dreams.

In another recent study (Coyle & Liben, 2015) we observed preschool girls and boys while they were playing with *GoldieBlox and the Spinning Machine* or with a masculinized version of the toy that we created (*BobbyBlox and the Spinning Machine*). Among the intriguing findings is that children mastered the belt-drive concept significantly better if they had been assigned to the toy of the other (i.e., mismatched) gender. Toy innovations and public relations campaigns may not always have their intended or expected effects. Although it would be premature to overextend the meaning of findings from laboratory studies like these, the findings do support the argument that it is crucial to investigate both intended and unintended consequences of interventions. It is particularly important to test whether interventions designed to attract girls and women to traditionally masculine domains serve to exacerbate constraining gender stereotypes and behaviors.

From the perspective of considering historical change and persistence of gendered phenomena, it is also interesting to compare the old advertising campaign designed by Philip Morris to attract women to Virginia Slims with the recent campaigns designed by the Science Cheerleaders and the European Commission to attract women to science. The substantive focus has unquestionably changed (science rather than smoking), but the medium has not (hyperfeminized women in both). Perhaps the medium is still the message. At the most general level, the lesson to be learned is the value of using theory and research in developmental science as a way to understand, influence, and evaluate decisions about how embedding layers of the ecological system should be structured. In the following section, I consider some key issues in linking gender-relevant developmental science to policy by focusing on one domain—the gender composition of schools and classrooms.

Single-Sex Schools

Unlike decisions about toy purchases that are made by individual parents and children, decisions about public-school structure are determined by federal and local policy makers. The legislative, executive, and judicial branches of the federal government formulate and enforce guiding laws and principles; local school boards and administrators select, implement, and monitor specific decisions about curriculum, textbooks, and the assignment of students to particular buildings and classrooms. The last of these decisions raises a multitude of contentious policy issues including the one I discuss here—the legitimacy of using gender as a basis for organizing public schools or classes. My three goals in discussing the debate about single-sex education are to demonstrate the practical relevance of theoretical and empirical work for gender-related policy, to extend earlier work on the values and challenges of connecting developmental science to practice (e.g., Huston, 2008; Shonkoff, 2000; Tseng, 2012), and to provide another illustration of the ways in which gender phenomena both change and remain similar over history.

Single-Sex Education in the United States

To provide a context for the contemporary debate about single-sex schooling, I begin with a very brief review of gender-differentiated education in the United States (see also Liben, 2015b). During the Colonial Period and the early years of the U.S. federacy, education for girls was focused largely on religious and moral instruction. Girls from privileged families received instruction in writing and arithmetic and sometimes in music and drawing, but not in the broader range of domains (e.g., geography, arithmetic) offered to boys. By the middle of the 19th century, both girls and boys received instruction in what were viewed as core disciplines (English, mathematics, arithmetic, and history), but instruction in some fields (e.g., science) remained gender specific. By the first half of the 20th century, most formal curricula were the same for boys and girls (common exceptions were physical education, health, and home economics), but many other educational influences (e.g., guidance counselors' course, college, and career recommendations) remained gender differentiated. During the late 1960s and early 1970s there was intensive political pressure for access to equivalent educational programs and institutions. In 1972, Title IX of the Civil Rights Act was passed, explicitly prohibiting discrimination based on sex in publicly funded education programs. This legislation is probably best known for its role in establishing gender equity in athletic programs, but it

also undergirded legal decisions that gave women access to previously all-male facilities and the inverse (e.g., opening the Virginia Military Institute to women and the nursing program of the Mississippi University for Women to men; see Brown, 2013; Sherwin, 2015).

Title IX was not, though, a panacea, and over the ensuing years, there were varied protests that public schools were still not serving girls and boys equitably. Among the criticisms were that girls were receiving less attention and fewer academic challenges from teachers, leading girls to develop more passive intellectual styles and avoid certain academic and career paths (e.g., Sadkar & Sadkar, 1995), and that boys were being alienated and discouraged by the highly feminized school environment, leading boys to evince more behavioral problems and higher school dropout (e.g., Hoff Sommers, 2000).

It was in this context that arguments for gender-differentiated education resurfaced. In 2006, the Department of Education issued revised regulations for Title IX, which permitted public schools to establish such programs under some conditions. In the ensuing decade, school districts increasingly instituted such programs in an effort to solve educational crises, particularly those in schools serving low-income and minority families. Although the absolute number of single-sex public school programs remains small in comparison to the roughly 13,000 school districts in the United States, and although precise numbers are difficult to obtain, estimates are that single-sex public schools or classes that numbered under a handful at the turn of the century now number close to a thousand (Klein, Lee, McKinsey, & Archer, 2014).

Given that research has demonstrated that gender stereotyping increases in response to heightened classroom attention to gender (e.g., Bigler, 1995; Hilliard & Liben, 2010), and given the lack of evidence for consistent academic benefits from single-sex environments, eight gender researchers (Rebecca Bigler, Lise Eliot, Richard Fabes, Diane Halpern, Laura Hanish, Janet Hyde, Carol Martin, and I) joined together to assemble and disseminate relevant scholarship to the wider community. Among our products was an article that appeared in the Educational Forum of the journal *Science* (Halpern et al., 2011). As might be inferred from its admittedly provocative title—*The Pseudoscience of Single-Sex Education*—the paper challenged the scientific basis for single-sex education. The paper attracted extensive media coverage (e.g., Chandler, 2011), and led to various attacks (e.g., Sax, 2011),

defenses (e.g., Fabes et al., 2011; Halpern et al., 2012), and debates (e.g., American Enterprise Institute, 2013). As a consequence of participating in or observing the ensuing exchanges, I have come to believe that developmental scientists' success in transferring research to policy will benefit from first, greater awareness of the sometimes competing interpretations of key constructs that underlie protagonists' positions; second, recognizing and respecting the range of evidentiary bases that hold sway in policy discussions; and third, greater attention to balancing the needs of individuals and of communities as a whole. Below I consider each of these in the context of the single-sex schooling debate.

How Is Gender Conceptualized?

One important issue to confront in the research-to-policy process is how diverse principals (i.e., researchers, practitioners, policy makers, and other stakeholders) define or conceptualize key constructs. In the case of single-sex education, this means examining foundational interpretations of the concept of gender itself. As argued in more detail elsewhere (see Liben, 2015b), my claim is that those on different sides of the single-sex education debate tend to hold differing foundational conceptualizations of gender, with most critics starting from a foundational commitment to gender constructivism and most proponents starting from a foundational commitment to gender essentialism. In the former, gender is conceptualized as emerging from the relational interplay of individuals and their embedding social context; in the latter, gender is conceptualized as inherent in males' and females' different underlying "essences." I have already sampled generously from the constructivist perspective in the prior sections of this article, and thus below I expand more fully on essentialist positions.

At the individual level, the notion of a gendered essence is linked to a range of genetic and hormonal processes; at the species level, it is linked to evolutionary processes (e.g., see discussions by Fine & Duke, 2015; Gelman & Taylor, 2000). These gendered essences are most often described as part of the "natural order," presumed to operate across eras, contexts, and portions of the life span. Deep-seated gendered essences may also be attributed to the work of a divine being (e.g., see exchange about benevolent sexism between Sax, 2002, and Glick & Fiske, 2002, discussed in Liben, 2015b). Whatever their origins, these essences are thought to have

varied and lasting consequences. In short, gender essentialism is a core belief that there are intrinsic, universal, natural, and pervasive differences between males and females across the life span, usually with the correlative belief that these differences are difficult to alter. Even if change can be achieved, the consequences are expected to be inauspicious.

A good illustration of a gender-essentialist perspective on education is provided by William DeWitt Hyde who was a highly influential educator early in the 20th century. He described pervasive differences between males' and females' life purposes and qualities. For example, in both economic and intellectual domains, W. Hyde (1906) characterized boys and men as producers but girls and women as consumers. About the economic domain he wrote:

. . . the manly economic ideal is the effective direction of production; the womanly ideal is the beneficent ordering of consumption. . . . Happy is the woman who as daughter, sister, wife, mother, finds herself excused from the task of direct economic production by the generous devotion of father, brother, husband, or son, and can find the economic justification of her life in this ministry and superintendence of the common household consumption. (W. Hyde, 1906, pp. 196–197)

With respect to the intellectual domain, W. Hyde suggested that boys' penchant for production enables men to become generative scientists, inventors, and businessmen as adults. In contrast, girls' penchant for consumption allows them to absorb already-established knowledge, which, in turn, positions them to transmit knowledge to new generations in adult roles as teachers. He claimed that these differences were rooted in nature, as when he stated that household functions are those "for which women are by nature and taste eminently fitted, and for which most manly men are conspicuously unfit" (p. 197).

W. Hyde (1906) further argued that because boys and girls are so different, "the methods of their education should be different. What is wholesome medicine for one is fatal poison for the other" (p. 205). He painted frightening pictures of the effects of ignoring the natural order, warning, for example, that schoolgirls "who allow themselves to be enveloped by intellectual ambitions" will be "guilty of a sin against the fountain-head of humanity, a crime against the race" (p. 203), and that:

Pitiful beyond expression is the mistake of those women who squander the wealth of physical vitality meant for twenty generations to gain some paltry academic honor or ephemeral social success. Terrible are the penalties nature exacts, — muscular flabbiness, nervous exhaustion, sharp-featured irritability, flat-chested sterility. (p. 203)

The flat-chested, sterile image of nontraditional, academically accomplished women painted by W. Hyde contrasts starkly to the decidedly not flat-chested image of nontraditional, scientifically accomplished women painted in the *Science Cheerleader* program or in the *Science: It's a Girl Thing!* campaign. In this sense, the messages about gender have changed dramatically from a century ago. In another sense, however, they have not: W. Hyde's arguments about the essential differences between boys and girls and about the pedagogical implications of these differences reappear in the writings of the two most visible and effective contemporary U.S. proponents of single-sex education—Michael Gurian and Leonard Sax. The major change is that the word *brain* now commonly appears in lieu of the word *nature*.

Illustratively, the first chapter of the book entitled *Boys and Girls Learn Differently* (Gurian, Henley, & Trueman, 2001) includes a seven-page table summarizing "Brain Gender Differences" and their implications for different behaviors and learning styles. In lectures to educators and parents, Gurian displays brain scans to document fundamental differences in the ways that boys' and girls' brains operate, and suggests that attempts to teach children without recognizing such differences are "potentially harmful for everyone" (Patch, 2013). Sax (2005) likewise catalogs a wide array of fundamental, biological differences between girls and boys that lead him to recommend various gender-distinct pedagogies. Illustratively, arguing that boys respond to stress with arousal and with the desire to repeat the emotion, Sax (2006) recommends that teachers confront boys aggressively by moving in close to them and shouting, for example, "What's your answer, Mr. Jackson? Give it to me!" (p. 193). Arguing that girls, in contrast, respond to stress by freezing up, Sax recommends addressing them softly and gently, as in "Lisa, sweetie, it's time to open your book. Emily, darling, would you please sit down for me and join us for this exercise?" (p. 195).

Similar positions that paint gender differences as natural (biological) are found in other contempo-

rary contexts as well. For example, in a trade book discussing gender from an evolutionary perspective, Gilbert (2006) observes that society has "ceased preparing boys and girls to be the different people they are" and laments that "Against every impetus of nature, we are *de*specializing the sexes" (p. 194). In a blog criticizing school antibullying programs as liberals' educational and political demasculinization of boys, McInnes (2013, paragraph 3) writes, "The fight against bullying is a great way to eradicate masculinity." In a debate on single-sex schooling, author Hoff Sommers scoffs researchers' recommendation that teachers avoid gender-specific language in classrooms, arguing that the gender distinction "is *not* something arbitrary and invidious that most of us *want* to leave behind" (American Enterprise Institute, 2013).

In summary and at the most general level, my argument is that the initial step in policy debates requires identification and discussion of alternative foundational concepts and values. Although such discussions may not always lead to agreement, they can at least identify where to direct additional conversation and debate. Without such efforts, and to adapt W. Hyde's turn of phrase, policy strategies that are seen as wholesome medicine by some will be seen as fatal poison by others.

What Counts as Evidence?

A second issue relevant to the science-to-policy process that emerges in the debate on single-sex schooling concerns the evidentiary bases for policy decisions, here evidence about the relative effectiveness of single-sex versus coeducational environments. Most developmental scientists are likely to assume that evidence must come from empirical studies that employ the strongest design, analytic, and interpretative tools of their science. However, scientists do not control the definition of evidence. In writing about this issue, Solesbury (2001), for example, commented:

> It is easy, especially for researchers, to elide the concepts of evidence and research, and for academic researchers to believe that only academic research counts [and that] somehow academic research is the basis of all knowledge. . . .In practice, evidence is more plural than research. (p. 8)

Other kinds of evidence that count in policy deliberations include anecdotes, expert judgment, testimony by individuals or groups, case narratives, and first-hand experiences of stakeholders and policy makers themselves. It is not that academic researchers necessarily shun these kinds of evidence (witness my own use of anecdotes in the current article), nor that practitioners rely on only their first-hand observations. But different kinds of evidence hold different sway in different disciplinary and experiential traditions.

Furthermore, it is not only that scientific evidence may be given less weight by nonacademic researchers; it is also that science may be actively rejected. Consider, for example, a comment posted by "bryan3" in response to a *Washington Post* newspaper article (Chandler, 2011) about the "Pseudoscience of Single-Sex Schooling" paper described earlier:

> The world is really going to hell. I mean. . . REALLY GOING TO HELL. When times were simpler, no one cared about "studies" that claimed "this is bad. . . that is bad. . . it's all bad. . . ." We just enjoyed life. Part of that was being able to go to an all boys school. . .or an all girls school. We never even imagined there should be "studies" claiming this was bad. I honestly pity you folks who will live in a world of "studies" in the future where you will be so scared to make a move, for fear that it will be a wrong one. How sad that you have missed out on a much simpler and infinitely better world where such BS wasn't even a thought.

Furthermore, even among those who agree that "studies" are important, individuals may view those studies through different lenses. For example, does the consumer judge a particular study in isolation or judge it within the context of the available corpus of relevant studies? Ask if crucial confounding variables were controlled or not? Consider if the sample was representative or not? Calculate whether the study was adequately powered or not? Explore whether the findings were replicated or not?

In the heat of policy debates (e.g., contentious school-board hearings) or mass-media venues (e.g., news articles or radio call-in shows) it is easy to cite isolated findings from individual studies that may be of questionable validity, replicability, or generalizability. Likewise, it is easy to offer only vague generalizations that seem to summarize a body of work, as in a recent news article about the decision to close an all-women's college that reported only that "Studies show that boys and girls test better and learn more in single-sex set-

tings" (Dvorak, 2015). Systematic reviews do not necessarily solve the problem given that there are also debates about which ones are sound.

Among proponents of single-sex education, a commonly cited review is one by Mael, Alonso, Gibson, Rogers, and Smith (2005) that had been commissioned by the U.S. Department of Education. Based on a vote-counting narrative method, Mael et al. concluded that although findings across individual studies were mixed, overall there was some advantage for single-sex programs. Later investigators, however, have criticized the sampling and analytic approaches used by Mael et al. Using sophisticated propensity matching and meta-analytic procedures, neither Nagengast, Marsh, and Hau (2013) nor Signorella, Hayes, and Li (2013) found an overall advantage for single-sex schools. In a still more recent meta-analysis, Pahlke, J. Hyde, and Allison (2014) confronted, head-on, the issue of the scientific quality of the studies by categorizing studies as controlled versus uncontrolled. Among uncontrolled studies, the data showed modest achievement advantages for single-sex programs. Among studies that controlled for selection effects, however, advantages for single-sex programs were trivial or entirely absent. Indeed, they reported that for some outcomes, the direction of the effect was reversed such that the advantage that did appear was for coeducational rather than for single-sex schools.

The response to the Pahlke et al. (2014) meta-analysis is also instructive. A few days after it was posted on the online platform of *Psychological Bulletin*, the paper received media attention (e.g., Preidt, 2014). A little over a week later, single-sex education advocate Michael Gurian criticized the paper on his *Single-Sex Education Blog*, disparaging both its online format and its meta-analytic approach. He wrote, "online releases are fine as opinion pieces, not studies; researchers do need to actually study children, in a rigorous and hands-on way, if they are going to claim they know what children need" (Gurian, 2014, point six, asterisk 4). It is unlikely that any academic researcher would disparage a meta-analysis as failing to represent actual studies or would presume an electronic preview of a *Psychological Bulletin* article to be of dubious rigor.

The issues just discussed suggest the need to target at least some dissemination and outreach efforts toward educating the lay public and policy makers in the values and ways of good science in general. I recognize that success will not be easy in a country in which policies related to global warming, evolu-

tion, and vaccination are also insufficiently influenced by science, but simply pounding away at disseminating scientific findings about a particular issue is not likely to be fully effective either. I should add here that although I have focused largely on the challenges of working at the intersection of the academy and the nonacademic world, there are also satisfactions. For example, developmental scientists' conceptual and empirical work on the consequences of gender segregation contributed to the work of the American Civil Liberties Union as it has documented and successfully challenged school districts' unconstitutional use of essentialist justifications for public single-sex programs (e.g., see American Civil Liberties Union, 2013, 2014; Sherwin, 2015).

How Are Individual and Community Needs Balanced?

The debate about single-sex education provides a rich domain in which to illustrate insights, challenges, and satisfactions in the science-to-policy process. I have focused on the importance of recognizing and confronting diversity in how domain-relevant constructs (like gender) are conceptualized and in how evidentiary sources (like meta-analyses and personal experiences) are understood and used. This debate also makes apparent the need to consider how a given policy simultaneously affects individuals and the community as a whole.

Selection processes may lead to systematic differences between students who are enrolled in different programs. For example, when policy makers provide optional school programs such as single-sex academies, the families who pursue them tend to have better resources than the families who do not. This factor, combined with other nonrandom selection forces (e.g., school admission policies, selective student withdrawal, and selective school dismissal) means that students who end up in single-sex programs differ from those in coeducational settings, in turn explaining why statistical controls for student variables are so essential for evaluating effects of school structure per se (Bigler, Hayes, & Liben, 2014; Hayes, Pahlke, & Bigler, 2011; Pahlke et al., 2014).

However, entry qualities are important not only as pretest data for examining the progress of individual students. They are also important because they affect the community as a whole. Individual students achieve more when they are surrounded by academically oriented and less disruptive peers, both of which have been linked to all girls' schools (see, respectively, Hayes et al.,

2011; Lavy & Schlosser, 2011). Thus, while individual students who are moved to a more selective environment may benefit from the move, it is also important to recognize that there may be a cost to the community that remains behind. Similar issues have been addressed with respect to the charter school movement and the consequences of school choice more generally (Chaltain, 2013; Liben, 2015b; Orfield & Frankenberg, 2013; Ravitch, 2013).

The balancing of individual and communal needs is also at issue when responding to breaches of traditional gender norms. Consider, for example, a news report of a 9-year-old boy who was bullied when he wore a Little Pony backpack to his (public) school, and the school responded by asking the family to stop their son from wearing the backpack to school (Cox, 2014). Or consider a report of a case in which the guardians of an 8-year-old girl were told that she would not be welcome to remain at her (private, religious) school unless she were to "dress and act more like a girl" (Bever, 2014). Although brief media reports like these cannot provide full and balanced detail, they nevertheless offer useful occasions to recognize tensions between individual and collective needs. In cases of gender-nonconforming students, should the school's goal be to change the individual child to fit the school ecology, or should the goal be to change the school ecology to accommodate the child? The view that traditional gender distinctions are desirable would motivate actions to reform (or remove) the gender-nontraditional student, whereas the view that traditional gender distinctions are unnecessary and constraining would motivate actions to reform the social ecology, perhaps by implementing an antisexism intervention such as the one described earlier (Lamb et al., 2009). In short, needs of individuals and of communities may not always be aligned, and different foundational values may lead to different policy choices.

The issues that have been raised in this section on the single-sex education debate are varied, and the discussions have necessarily been abbreviated (for additional detail, see Liben, 2015b; Signorella & Bigler, 2011, 2013). All instantiate a premise with which I began this article—that human development can be understood only by considering the individual within the embedding ecology, or more broadly, as part of a relational developmental system (Overton, 2013). In my closing comments I reflect on the importance of recognizing that constructive and socialization processes operate within an embedding societal context that itself changes over time.

Looking Back and Looking Ahead

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started,
And know the place for the first time.

T. S. Eliot (1942)

The claim in the title of this article that "we've come a long way" is supported by reflecting on earlier times. By today's standards, it seems shocking to be confronted by an era in which there was a 10% quota on women medical students, a lack of sleeping accommodations for women interns, explicit gender-based discouragement of a young girl's interest in pursuing a legal career coupled with the requirement that she compete to become the Betty Crocker Homemaker of Tomorrow. There have been striking changes in domains like these, with documented improvements in gender parity for a range of educational, occupational, and personal experiences and outcomes.

At the same time, the claim that "we're not there yet" also finds convincing support. It is not only that disparities continue in these same domains; it is also that many of the observations that appear to be so shocking when viewed in their historical contexts have surprisingly close analogs in contemporary times. The arguments about the bases for gender-differentiated education that were made over a century ago by W. Hyde have surprisingly close analogs in books by Sax and Gurian written within the last decade; the sexual tone of the advertising campaign used for Virginia Slims in the 1960s is echoed in campaigns designed 50 years later to attract more girls and women to science.

These and additional examples from personal and societal history suggest that even as there have been significant and successful efforts to reduce gender-based constraints on developmental outcomes, there remain powerful forces operating to sustain them. One conservative force is the constructive nature of individual children's information processing. Children actively build gender schemata about what is differentially appropriate for boys and for girls, and then use these schemata as guides to selectively engage with, interpret, and remember what is available in the embedding context. Empirical work has shown, for example, that children have more difficulty recalling material that is inconsistent (rather than consistent) with cultural gender stereotypes. Likewise, children are less likely to participate in activities viewed as appropriate for the

other gender. Individually driven processes like these thus continue to steer girls' and boys' experiences, aspirations, and outcomes along somewhat different paths.

Also having a conservative effect are the socializing agents—parents, teachers, and peers—who continue to encourage, model, and enforce traditional gender messages. Sometimes these messages are conveyed unintentionally as when teachers who are committed to enacting gender-neutral strategies nevertheless routinely comment on young girls' appearance and on young boys' strength. But sometimes gender-traditional messages are conveyed deliberately because traditional gender differences are explicitly valued. Many fathers and mothers dream of raising their sons and daughters to join them in traditional masculine and feminine pastimes and to carry on gendered traditions; some educators and authors argue explicitly that it is a societal mistake to despecialize the sexes (Gilbert, 2006), eradicate masculine behaviors like bullying (McInnes, 2013), or leave gender distinctions behind (American Enterprise Institute, 2013; Hoff Sommers, 2000).

Before concluding that there is largely historical stability, however, it is important to recognize that far more radical revisions about gender concepts and use have been emerging in the very recent past. One such change is a shift away from the traditional gender binary. Illustratively, rather than just "male" and "female," Facebook now allows users to select up to 10 of 50 gender options for gender self-identification (Oremus, 2014; Weber, 2014). Another label-related change is the current movement to replace gendered language with gender-neutral language (e.g., "ze" in place of he and she; "hir" in place of his or hers; Bigler & Leaper, 2015). Also breaking with tradition is the growing rejection of the assumption that gender is directly and permanently linked to birth sex. Gender identities are increasingly understood as reflections of one's personal sense of self as male or female rather than as reflections of one's biological qualities. Furthermore, individuals are increasingly able to modify those qualities through medical interventions. There is also growing acceptance that neither gender nor birth sex automatically defines social roles or relationships, for example, defining who can join together as romantic couples or parenting pairs. Educational institutions are designing and instituting innovative policies and programs to address these changing concepts of gender (Mooney, 2015).

Taken together, these various changes mean that gender and physical appearance may or may not align in historically traditional ways, may or may not be marked in language, and may or may not be permanent over the life span or even across settings. Such conditions can be expected to make both gender and biological qualities increasingly complex and fluid, thereby making it increasingly more difficult to assign people into one of two dichotomous gender categories for any purpose, be it classroom instruction, bathrooms, locker rooms, or sports teams. Existing theories—for example, DIT—would predict very different developmental outcomes if the current embedding context that routinely labels and separates visibly distinct males and females were replaced by an embedding context that groups people not by gender, but rather by criteria that change in accord with the particular purpose at hand (e.g., using height to assign players to basketball leagues and hiking skills to assign youth to scouting troops).

In closing, I again return to the claim conveyed by the article's title, that although we have come a long way, we are not there yet. I have discussed ways in which gender distinctions, concepts, and constraints have and have not changed over roughly the last century. But where, exactly, is the "there" that we want to reach? As I hope is apparent from my earlier descriptions of alternative values and conceptualizations of gender, there is unlikely to be consensus about the answer to that question. But the likely cacophony of responses does not mean that we should simply ignore the question. Indeed, I would argue that as developmental scientists, it is essential that we grapple with it. As Cassirer (1944) wrote about development:

> We cannot describe the momentary state of an organism without taking its history into consideration, and without referring it to a future state for which this state is merely a point of passage. (p. 50)

SRCD's twofold mission—to advance developmental science and promote its use to improve human lives—requires that we think deeply and respectfully both about our history and about our future as we study human development in our present.

References

Allport, G. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.

Amazon. (2015). LEGO Pink Brick Box (5585). Retrieved from http://www.amazon.com/dp/B000Y8EPHW/ref=pdp_new_dp_review

American Civil Liberties Union. (2013, July 8). Court approves settlement reached in challenge to West Virginia single-sex school program rooted in stereotypes. Retrieved from https://www.aclu.org/news/court-approves-settlement-reached-challenge-west-virginia-single-sex-school-program-rooted

American Civil Liberties Union. (2014, December 1). Teach kids, not stereotypes. Retrieved from https://www.aclu.org/womens-rights/teach-kids-not-stereotypes

American Enterprise Institute. (2013, August 28). Should single-sex schooling be eliminated? Debate between Christina Hoff Sommers and Lise Eliot. Retrieved from http://www.aei.org/events/should-single-sex-schooling-be-eliminated/

Association of American Medical Colleges. (2014). *Physician specialty data book 2014*. Center for Workforce Studies. Retrieved from https://www.aamc.org/data/facts/

Association of American Medical Colleges. (2015). FACTS: Applicants, matriculates, enrollment, graduates, MD/PhD, and residency applicants data. Table 28. Retrieved from https://www.aamc.org/data/facts/

Bever, L. (2014, March 26). Virginia girl's style not feminine enough for Lynchburg Christian School. *The Washington Post*. Retrieved from http://www.washingtonpost.com/news/morning-mix/wp/2014/03/26/virginia-girls-toys-not-feminine-enough-for-lynchburg-christian-school/

Bigler, R. S. (1995). The role of classification skill in moderating environmental influences on children's gender stereotyping: A study of the functional use of gender in the classroom. *Child Development*, *66*, 1072–1087. doi:10.2307/1131799

Bigler, R. S. (2013). Understanding and reducing social stereotyping and prejudice among children. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world* (pp. 327–331). Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780199890712.003.0060

Bigler, R. S., Brown, C. S., & Markell, M. (2001). When groups are not created equal: Effects of group status on the formation of intergroup attitudes in children. *Child Development*, *72*, 1151–1162. doi:10.1111/1467-8624.00339

Bigler, R. S., Hayes, A. R., & Liben, L. S. (2014). Analysis and evaluation of the rationales for single-sex schooling. In L. S. Liben & R. S. Bigler (Vol. Eds.) & J. Benson (Series Ed.), *The role of gender in educational contexts and outcomes: Vol. 47. Advances in child development and behavior* (pp. 225–260). San Diego, CA: Elsevier. doi:10.1016/bs.acdb.2014.05.002

Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development*, *68*, 530–543. doi:10.2307/1131676

Bigler, R. S., & Leaper, C. (2015). Gendered language: Psychological principles, evolving practices, and inclusive policies. *Policy Insights from the Behavioral and Brain Sciences*, *2*, 187–194. doi:10.1177/2372732215600452

Bigler, R. S., & Liben, L. S. (1990). The role of attitudes and interventions in gender-schematic processing. *Child Development*, *61*, 1440–1452. doi:10.2307/1130754

Bigler, R. S., & Liben, L. S. (1992). Cognitive mechanisms in children's gender stereotyping: Theoretical and educational implications of a cognitive-based intervention. *Child Development*, *63*, 1351–1363. doi:10.2307/1131561

Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 34, pp. 39–89). San Diego, CA: Elsevier. doi:10.1016/S0065-2407(06)80004-2

Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science*, *16*, 162–166. doi:10.1111/j.1467-8721.2007.00496.x

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

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*, 513–531. doi:10.1037/0003-066X.32.7.513

Brown, C. S. (2013). Legal issues surrounding single-sex schools in the U.S.: Trends, court cases, and conflicting laws. *Sex Roles*, *69*, 356–362. doi:10.1007/s11199-011-0001-x

Cassirer, E. (1944). *An essay on man: An introduction to a philosophy of human culture*. New Haven, CT: Yale University Press.

Cavalier, D. (2015). *Science cheerleader*. Retrieved from www.sciencecheerleader.com

Chaltain, S. (2013, October 16). The moral limits of school choice. Of, by, for: In search of the civic mission of K-12 schools. *Education Week*. Retrieved from http://blogs.edweek.org/edweek/civic_mission/2013/10/the_moral_limits_of_school_choice.html

Chandler, M. A. (2011, September 22). Study: Single-sex education may do more harm than good. *The Washington Post*. Retrieved from http://www.washingtonpost.com/local/education/study-single-sex-education-may-do-more-harm-than-good/2011/09/22/gIQABAQOoK_story.html?hpid=z4

Collins, N. (2012, June 22). Science 'girl thing' video branded offensive. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/news/science/science-news/9349923/Science-girl-thing-video-branded-offensive.html

Cox, L. (2014, March 18). 'My Little Pony' backpack banned—School says young boy will be bullied. *Hollywood Life*. Retrieved from http://hollywoodlife.com/2014/03/18/my-little-pony-backpack-banned-school-bullies/

Coyle, E. F., & Liben, L. S. (2015, March). *Preschoolers' mechanical learning from a STEM toy: The roles of parental guidance and gender-targeted marketing*. Poster presented

at the biennial meetings of the Society for Research in Child Development, Philadelphia, PA.

Coyle, E. F., & Liben, L. S. (in press). Affecting girls' job and activity interests through play: The moderating roles of personal gender salience and game characteristics. *Child Development*. Advance online publication. doi:10.1111/cdev.12463

Dvorak, P. (2015, March 7). Sweet Briar College closing sours the notion of equality for women. *The Washington Post*. Retrieved from https://www.washingtonpost.com/local/sweet-briar-college-closing-sours-notion-of-equality-for-women/2015/03/07/dd4a4588-c4d7-11e4-ad5c-3b8ce89f1b89_story.html

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 paractice and policy* (pp. 19–34). New York, NY: Oxford University Press.

Elder, G. H., Jr. (1998). The life course and human development. In R. M. Lerner (Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (5th ed., pp. 939–991). New York, NY: Wiley.

Eliot, T. S. (1942). *The little gidding*. London, UK: Faber and Faber.

European Commission. (2015). *Science: It's a girl thing!* Women in research and education. An initiative of the European Commission. Retrieved from http://science-girl-thing.eu/en

Fabes, R. A., Martin, C. L., Hanish, L. D., Bigler, R. S., Eliot, L., Halpern, D. F., Hyde, J., & Liben, L. S. (2011, September 28). A reply to Sax on the dangers of single-sex schools [Web log post]. *Psychology Today Blogs in The Guest Room*. Retrieved from https://www.psychologytoday.com/blog/the-guest-room/201109/reply-sax-the-dangers-single-sex-schools

Fine, C., & Duke, R. (2015). Expanding the role of gender essentialism in the single-sex education debate: A commentary on liben. *Sex Roles*, *72*, 427–433. doi:10.1007/s11199-015-0474-0

Gelman, S. A., & Taylor, M. G. (2000). Gender essentialism in cognitive development. In P. H. Miller & E. K. Scholnick (Eds.), *Toward a feminist developmental psychology* (pp. 169–190). New York, NY: Routledge.

Gilbert, M. (2006). *The disposable male: Sex, love, and money—Your world through Darwin's eyes*. Oversheil Farm, UK: Hunter Press.

Glick, P., & Fiske, S. T. (2002). Ambivalent responses. *American Psychologist*, *56*, 444–446. doi:10.1037/0003-066X.57.6-7.444b

Gurian, M. (2014, February 18). Single-sex education challenge. *Gurian Institute*. Retrieved from http://gurianinstitute.com/wordpress/single-sex-education-challenge/

Gurian, M., Henley, P., & Trueman, T. (2001). *Boys and girls learn differently! A guide for teachers and parents*. San Francisco, CA: Jossey-Bass.

Halpern, D. F., Eliot, L., Bigler, R. S., Fabes, R. A., Hanish, L. D., Hyde, J., Liben, L. S., & Martin, C. L. (2011).

The pseudoscience of single-sex schooling. *Science*, *333*, 1706–1707. doi:10.1126/science.1205031

Halpern, D. F., Eliot, L., Bigler, R. S., Fabes, R. A., Hanish, L. D., Hyde, J., Liben, L. S., & Martin, C. L. (2012). Response to "The pseudoscience of single-sex schooling." *Science*, *335*, 166–168. doi:0.1126/science.1205031

Hayes, A. R., Pahlke, E. E., & Bigler, R. S. (2011). The efficacy of single-sex education: Testing for selection and peer quality effects. *Sex Roles*, *65*, 693–703. doi:10.1007/s11199-010-9903-2

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

Hines, M. (2015). Gendered development. In M. E. Lamb (Ed.), *Handbook of child development and developmental science: Vol. 3. Socioemotional processes* (7th ed., pp. 842–887). Hoboken, NJ: Wiley.

Hoff Sommers, C. (2000). *The war against boys: How misguided feminism is harming our young men*. New York, NY: Simon & Shuster.

Huston, A. C. (2008). From research to policy and back. *Child Development*, *79*, 1–12. doi:10.1111/j.1467-8624.2007.01107.x

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*, 581–592. doi:10.1037/0003-066X.60.6.581

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, *104*, 53–69. doi:10.1037/0033-2909.104.1.53

Hyde, W. D. (1906). *The college man and the college woman*. Boston, MA: Houghton Mifflin.

Klein, S., Lee, J., McKinsey, P., & Archer, C. (2014, December 11). Identifying US K-12 public schools with deliberate sex segregation. Feminist Majority Foundation. Retrieved from http://feminist.org/education/pdfs/IdentifyingSexSegregation12-12-14.pdf

Lamb, L. M., Bigler, R. S., Liben, L. S., & Green, V. A. (2009). Teaching children to confront peers' sexist remarks: Implications for theories of gender development and educational practice. *Sex Roles*, *61*, 361–382. doi:10.1007/s11199-009-9634-4

Lavy, V., & Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, *3*, 1–33. doi:10.1257/app.3.2.1

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

LEGO Shop. (2015). *LEGO Friends Butterfly Beauty Shop*. Retrieved from http://shop.lego.com/en-US/Butterfly-Beauty-Shop-3187

Liben, L. S. (1974). Operative understanding of horizontality and its relation to long–term memory. *Child Development*, *45*, 416–424.

Liben, L. S. (2010, April). Single sex schools: Out of the pan and into the fire. In. M. L. Signorella (Chair), *Single sex schools: Debates, data, and decision making*. Symposium presentation at the Gender Development Research Conference, San Francisco, CA.

Liben, L. S. (2014). Perceiving and representing horizontals: From laboratories to natural environments and back again. *Ecological Psychology*, 26, 158–166. doi:10.1080/10407413.2014.875341

Liben, L. S. (2015a). The STEM gender gap: The case for spatial interventions. *International Journal of Gender, Science, and Technology*, 7, 133–150.

Liben, L. S. (2015b). Probability values and human values in evaluating single-sex education. *Sex Roles*, 72, 401–426. doi:10.1007/s11199-014-0438-9

Liben, L. S., & Bigler, R. S. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development*, 67(2, Serial No. 269). doi:10.1111/1540-5834.00200

Liben, L. S., & Bigler, R. S. (2015). Understanding and undermining the development of gender dichotomies: The legacy of Sandra Lipsitz Bem. *Sex Roles*. Advance online publication. doi:10.1007/s11199-015-0519-4

Liben, L. S., Bigler, R. S., & Hilliard, L. J. (2014). Gender development: From universality to individuality. 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. 3–18). New York, NY: Oxford University Press.

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

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

Liben, L. S., & Signorella, M. L. (1993). Gender-schematic processing in children: The role of initial interpretations of stimuli. *Developmental Psychology*, 29, 141–149. doi:10.1037/0012-1649.29.1.141

Mael, F., Alonso, A., Gibson, D., Rogers, K., & Smith, M. (2005). Single-sex versus coeducational schooling: A systematic review (No. 2005–01). U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. Retrieved from http://www2.ed.gov/rschstat/eval/other/single-sex/single-sex.pdf

Martin, C. L., Eisenbud, L., & Rose, H. (1995). Children's gender-based reasoning about toys. *Child Development*, 66, 1453–1471. doi:10.2307/1131657

Martin, C. L., & Fabes, R. A. (2001). The stability and consequences of young children's same-sex peer interactions. *Developmental Psychology*, 37, 431–446. doi:10.1037//0012-1649.37.3.431

Martin, C. L., Fabes, R. A., & Hanish, L. (2014). Gendered-peer relationships in educational contexts. In L. S. Liben & R. S. Bigler (Vol. Eds.) & J. Benson (Series Ed.), *The role of gender in educational contexts and outcomes. Advances in child development and behavior* (pp. 151–187). San Diego, CA: Elsevier.

Martin, C. L., & Halverson, C. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development*, 52, 1119–1134. doi:10.2307/1129498

McInnes, G. (2013, May 17). Let's not let boys be boys. *Taki's Magazine*. Retrieved from http://takimag.com/article/lets_not_let_boys_be_boys_gavin_mcinnes/print#axzz2aLQbjylC

Mooney, C. (Ed.). (2015, October 23) Diversity in academe. *The Chronicle of Higher Education*. Section B.

Nagengast, B., Marsh, H. W., & Hau, K.-T. (2013). Effects of single-sex schooling in the final years of high school: A comparison of analysis of covariance and propensity score matching. *Sex Roles*, 69, 404–422. doi:10.1007/s11199-013-0261-8

Oremus, W. (2014, February 13). Here are all the different genders you can be on Facebook. *Slate*. Retrieved from http://www.slate.com/blogs/future_tense/2014/02/13/facebook_custom_gender_options_here_are_all_56_custom_options.html

Orfield, G., & Frankenberg, E. (Eds.). (2013). *Educational delusions? Why choice can deepen inequality and how to make schools fair*. Berkeley: University of California Press.

Overton, W. F. (2013). A new paradigm for developmental science: Relationism and relational-developmental-systems. *Applied Developmental Science*, 17, 94–107. doi:10.1080/10888691.2013.778717

Pahlke, E., Hyde, J. S., & Allison, C. M. (2014). The effects of single-sex compared with coeducational schooling on students' performance and attitudes: A meta-analysis. *Psychological Bulletin*, 140, 1042–1072. doi:10.1037/a0035740

Patch, D. (2013, February 8). Boys and girls don't learn the same way, author insists. *The Toledo Blade*. Retrieved from http://www.toledoblade.com/local/2013/02/08/boys-and-girls-don-t-learn-the-same-way-author-insists.html

Peacock, L. (2012, December 7). 'Sexy' science grads make spoof video to combat sexism. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/women/sex/9730203/Sexy-science-grads-make-spoof-video-to-combat-sexism.html

Pratini, N. (2014, May 6). Male doctors earn 88 percent more through Medicare, new study shows. *Huffington Post*. Retrieved from http://www.huffingtonpost.com/napala-pratini/male-doctors-earn-88-more_b_5267938.html

Preidt, R. (2014, February 7). Single-sex education's benefits challenged in study. *HealthDay*. Retrieved from http://consumer.healthday.com/kids-health-information-23/child-development-news-124/single-sex-education-s-benefits-challenged-in-study-684512.html

Ravitch, D. (2013). *Reign of error: The hoax of the privatization movement and the danger to America's public schools*. New York, NY: Knopf.

Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 858–932). Hoboken, NJ: Wiley.

Sadkar, M., & Sadkar, D. (1995). *Failing at fairness: How our schools cheat girls*. New York, NY: Touchstone.

Santayana, G. (1905). *The life of reason or the phases of human progress: Reason in common sense*. New York, NY: Scribner's.

Sax, L. (2002). Maybe men and women are different. *American Psychologist*, *57*, 444. doi:10.1037/0003-066X.57.6-7.444a

Sax, L. (2005). *Why gender matters*. New York, NY: Doubleday.

Sax, L. (2006). Six degrees of separation: What teachers need to know about the emerging science of sex differences. *Educational Horizons*, *84*, 190–200.

Sax, L. (2011, September 26). Are single-sex schools actually dangerous? Sax on Sex Blog [Web log post]. *Psychology Today*. Retrieved from http://www.psychologytoday.com/blog/sax-sex/201109/are-single-sex-schools-actually-dangerous

Sherwin, G. (2015). Anecdotal and essentialist arguments for single-sex educational programs discussed by Liben: A legal analysis. *Sex Roles*, *72*, 434–445. doi:10.1007/s11199-015-0478-9

Shonkoff, J. P. (2000). Science, policy and practice: Three cultures in search of a shared mission. *Child Development*, *71*, 181–187. doi:10.1111/1467-8624.00132

Signorella, M. L., & Bigler, R. S. (2011). Single-sex schooling: Part I [Special issue]. *Sex Roles*, *65*.

Signorella, M. L., & Bigler, R. S. (2013). Single-sex schooling: Part II [Special issue]. *Sex Roles*, *69*.

Signorella, M. L., Hayes, A. R., & Li, Y. (2013). A meta-analytic critique of Mael et al.'s (2005) review of single-sex schooling. *Sex Roles*, *69*, 423–441. doi:10.1007/s11199-013-0288-x

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. doi:10.2307/1129951

Signorella, M. L., & Liben, L. S. (1985). Assessing children's gender–related attitudes. *Psychological Documents*, *15*, 7.

Society for Research in Child Development. (2015). *About us: Mission and scientific vision*. Retrieved from http://srcd.org/about-us/strategic-plan/mission-and-scientific-vision

Solesbury, W. (2001). Evidence based policy: Whence it came and where it's going. ESRC UK Center for Evidence Based Policy and Practice: Working Paper I. London, UK: University of London.

Sterling, D. (2013, March 17). Inspiring the next generation of female engineers: Debbie Sterling at TEDxPSU. Retrieved from http://tedxtalks.ted.com/video/Inspiring-the-next-generation-2

Sterling, D. (2015). I'm an engineer: Debbie Sterling. Retrieved from http://www.engineergirl.org/Engineers/Directory/13512.aspx

The Young Turks. (2012, July 2). Science: It's a girl thing! *The Young Turks*. Retrieved from https://www.youtube.com/watch?v=Y8UdvBJUZ9U

Tseng, V. (2012). The uses of research in policy and practice. *Social Policy Report of the Society for Research in Child Development*, *26*(2), 1–16.

Virginia Slims Commercials. (1969, September). Compilation of Virginia Slims cigarette television commercials including "You've Come a Long Way Baby." Retrieved from: https://industrydocuments.library.ucsf.edu/tobacco/docs/yhyd0111

Weber, P. (2014, February 21). Confused by all the new Facebook genders? Here's what they mean. *TheWeek.com*. Retrieved from http://www.slate.com/blogs/lexicon_valley/2014/02/21/gender_facebook_now_has_56_categories_to_choose_from_including_cisgender.html