Gender-Differentiating Effects of Parenthood on Income among High-Potential Individuals in the Context of Other Valued Aspects of Life

By

Brian Bernstein

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Psychological Sciences

October 31st, 2021

Nashville, Tennessee

Approved:

Professor David Lubinski

Professor Camilla P. Benbow

Professor Sonya K. Sterba

Professor Andrew J. Tomarken

To my parents, Robert and Nancy,

who gave me the opportunity to pursue my own unique interests and abilities.

ACKNOWLEDGEMENTS

This research was funded by the National Science Foundation Graduate Research Fellowship, the Templeton Foundation (#61368), and the Eunice Kennedy Shriver National Institute of Child Health and Human Development under Award U54HD083211.

I was fortunate in that I was able to find a home in the international and Jewish communities in Nashville during my time in the city. I also benefitted tremendously from an intellectually engaging group of peers on Vanderbilt's campus. The faculty in my program area regularly challenged and supported me as a critical thinker and analyst.

Rachel Rentfro, Kira McCabe, and Camilla Benbow consistently encouraged me and helped me to develop. It was a pleasure to work with them, and I became a better scholar through my collaboration with them.

I am honored by the investments that these people made in my development, but I would be remiss if I did not mention the person with whom I worked most closely over the last five years: David Lubinski. It is difficult to find true advocates in life—someone who is in your corner of the ring no matter the circumstances and no matter the time and energy it requires. I was lucky to find such a person in David.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE GENDER WAGE GAP

Among full-time, year-round workers, women in the United States earn 81.5 cents for every dollar earned by men (Bureau of Labor Statistics, 2020; Hegewisch & Williams-Baron, 2018). Given the magnitude of this difference, this disparity is found to be deeply unsettling and has sparked federal legislation, public conversation, and multiple lines of research throughout the social sciences dedicated to identifying the underlying causes of this gender disparity. Even a decade ago, this issue had received such attention that it was ranked seventh on *Nature's* top ten list of pressing questions facing social scientists: "Why do so many female workers still earn less than male workers?" (Giles, 2011, p. 18).

Investigations into income differences between women and men have led to the discovery of some interesting subtleties involving multiple determinants. Some of the determinants of this phenomenon can be traced to preferences for contrasting educational and occupational choices with different pay scales (Okahana & Zhou, 2018; Snyder et al., 2019); others involve individual differences in the many different ways in which people choose to structure their lives (Browne, 2002, 2005; Ferriman et al., 2009; Geary, 2021; Hakim, 2017; Lubinski et al., 2014; Susan Pinker, 2008; Rhoads, 2004). Still, what we have learned over the past decade does not fully explain this gender difference. The objective of this dissertation is to advance the understanding of the effects of parenthood on the gender wage gap in high-potential men and women. Moreover, in addition to examining how and why gender differences in income emerge, I will

assess in parallel over the life course how other personally valued outcomes vary among women and men. That is, a more holistic approach to the gender wage gap will be implemented to place this phenomenon in a broader context by examining individual and group differences in what constitutes a meaningful and satisfying life and the extent to which different life paths covary with psychological well-being. Just as Kahneman and Tversky (Kahneman, 2011; Kahneman et al., 1982; Kahneman & Tversky, 2000) showed that there is more to economic decision making than maximizing profit (because competing preferences are always at play), so too is there more to constructing a meaningful and satisfying life than maximizing income (because there are individual differences in career and lifestyle preferences). Given that other valued life outcomes are seldom assessed concurrently with dynamic income changes among the genders following parenthood, this dissertation will attempt to fill this gap.

In Chapter 1 of this dissertation, I will first review the dimensions known to give rise to the gender wage gap. Then, I will review some determinants that have changed over the past two decades as well as others that have remained in place. In the process, I will examine some especially fruitful areas in need of research. Finally, these considerations will be connected to the specific purpose and scope of the empirical advances I hope to contribute by examining the gender wage gap in the broader context of lifespan development.

In Chapter 2, I will introduce a series of research hypotheses concerning the impact of parenthood on important life outcomes in high-potential populations, as well as the methodology used to address these hypotheses. Two sets of results and discussion from two samples of high-potential individuals (Study 1 and Study 2) are then given to address these hypotheses. Finally, general discussion and interpretation of the impact of parenthood on high-potential populations—and the relationship to the broader gender wage gap—is given.

Nuances of the Gender Wage Gap

The gender wage gap has not been constant across time; in fact, the gap is now smaller than it was in 1980 when comparable data were first collected. At that time, women earned 60.2% of what men earned (Bureau of Labor Statistics, 2019). While several hypothesized causes for the gender wage gap still operate today (e.g., discrimination, differential hands-on investments by parents in the home and family), other important determinants of a wage differential between the genders have been removed. Most notably, women in the 21st century are earning advanced degrees and credentials at rates greater than men, a cross-cultural phenomenon in highly developed countries (Stoet & Geary, 2020). For example, consistently over the past 15 years more women than men in the U.S. have been awarded doctorates (Okahana & Zhou, 2018; Snyder et al., 2019), and this reversal in credentialing has lessened the gender wage gap (Blau & Kahn, 2017).

To the extent that women and men have reached parity in postsecondary degrees and professional credentials, and to the extent that these accomplishments are related to occupational income, gender parity in income might be anticipated. However, gender parity in earnings has not been reached.

Additionally, it is important to note that, just as the gender wage gap has not been constant across time, it is also not constant across occupations. Influential work by Claudia Goldin (2014) has shown that careers with more flexible hours have the narrowest wage gaps (e.g., pharmacy; Goldin & Katz, 2016); meanwhile, those careers that are most demanding and allow the least amount of flexibility in work schedules, or are characterized by an unpredictable work schedules (e.g., lawyers and MBAs in major firms), show the widest wage gaps. Especially

within the most demanding and inflexible careers, gender differences in willingness to take on long and inflexible schedules lead to differences in pay (E. Becker & Lindsay, 2004; J. C. Williams & Boushey, 2010). These considerations underlying the gap have suggested the importance of studying incumbents in the most demanding and highest-pressure occupations, which also tend to be the least flexible. It is expected that, consistent with Goldin's work, women in these careers would see the largest wage gaps relative to their male colleagues. For example, in many of these occupations, compensation is a function of willingness to be on call for weekends and/or travel; to the extent that there are gender differences in a willingness to do so, income differences logically follow.

Finally, it is important to note that the gender wage gap changes depending on how one defines it. Specifically, the gender wage gap can be defined as an adjusted or non-adjusted gap. This distinction is important. In the adjusted wage gap, factors such as hours worked, level and area of educational achievement, and occupations chosen are controlled for. In studies that control for these determinants, women earn as much as 95 cents per dollar earned by men (Blau & Kahn, 2017; Chamberlain et al., 2019; Corbett & Hill, 2012; U.S. Department of Labor, 2009). A marked reduction occurs from the non-adjusted wage gap statistic of women earning 81.5 cents per dollar relative to men (Bureau of Labor Statistics, 2020; Hegewisch & Williams-Baron, 2018). This variable nature of the focal gender difference, as a function of the way in which groups are selected (or matched by covariates), has led to the acknowledgement of the importance of specific, well-defined groups when studying gender differentials in wages (Browne, 2018; Goldin, 2014).

Well over half of the discrepancy between men and women can be accounted for by selecting carefully specific subgroups or matching on relevant gender-differentiating

determinants. Nevertheless, many studies in the social sciences are launched, analyzed, and interpreted without taking hours worked, level and area of educational attainment, and occupational self-segregation into account. These dimensions all manifest meaningful gender differences over the course of development (Geary, 2021; Gino et al., 2015; Hakim, 2017; Stoet & Geary, 2018). When one does take these factors into account, the income disparity shrinks appreciably. Nevertheless, even with these determinants factored in, at least 5% of the total variance is still at play, and this amount of variance can have huge implications (Abelson, 1985; Kuncel & Hezlett, 2010; Lubinski & Humphreys, 1996; Taylor & Russell, 1939). Before examining possible causes in the remaining variance in the gender wage gap, however, it will be useful to detail some known determinants of educational/occupational choice and performance, which in part explain differential educational and occupational choices that result in careers that pay differently.

Causes of the Gender Wage Gap

Educational and Occupational Choices

Following the Theory of Work Adjustment (Dawis & Lofquist, 1984; Lofquist & Dawis, 1991; Tinsley, 1993) and the accrued empirical evidence supporting its verisimilitude over multiple decades (Bernstein et al., 2019; Dawis, 1992; Lubinski, 2016; Sackett et al., 2017), abilities and interests are central determinants of educational and occupational choices as well as performance after choices are made. These determinants are key psychological attributes required for understanding outcome differences for individuals as well as groups. Group differences are simply aggregated individual differences. So, when outcome differences are

anticipated among individuals who vary on attributes giving rise to differential outcomes, and if gender differences are consistently observed on these determinants, outcome differences among the genders are anticipated as well (Lubinski, 2020).

For example, multiple centennial reviews on the study of individual differences have shown that both level and pattern of intellectual abilities and interests structure important educational and occupational decisions and performance in learning and work settings (Dawis, 1992; Lubinski, 2016; Sackett et al., 2017). While the genders do not differ in overall intellectual ability level (Hunt, 2011; Jensen, 1998), there are appreciable gender differences in ability and interest patterns. For instance, there is a one standard deviation difference (male minus female) in interests for things versus people; there is also a one standard deviation difference (male minus female) in spatial ability minus verbal ability tilt (Hedges & Nowell, 1995; Su et al., 2009; Wai et al., 2009). These differences lead to the overrepresentation of women in humanistic/organic disciplines and their underrepresentation in STEM fields; the inverse is true for males. Careers in STEM are generally more lucrative, relative to the majority of career tracks. In addition, the greater variability of males, relative to females, in mathematical and spatial reasoning abilities results in greater male to female ratios at both distributional extremes (Arden & Plomin, 2006; Hedges & Nowell, 1995; Humphreys, 1988; Stewart-Williams & Halsey, 2021; Wai et al., 2010). The upper tails of mathematical and spatial reasoning distributions is the region from which the most lucrative and demanding STEM opportunities select (Johnson et al., 2008; O'Dea et al., 2018; Steven Pinker, 2002). This dissertation is not about individual differences that give rise to contrasting educational and occupational paths, but they are an important set of considerations to take into account as one aspect of why gender disparities in income are observed in the world of work.

Other relevant determinants include gender differences in lifestyle preferences and the characteristics of work environments (Geary, 2021; Hakim, 2017; Susan Pinker, 2008; Rhoads, 2004). Some work environments are aesthetically unpleasant, noisy, unclean, dangerous and highly stressful. Because they are less attractive to prospective employees and fewer employees are willing to take these jobs, those who do are highly compensated (Susan Pinker, 2008). In addition, higher levels of risk-taking behavior in men (Byrnes et al., 1999; Charness & Gneezy, 2012; Geary, 2021, pp. 236-238) are helpful for leadership roles in entrepreneurship and high-growth companies as well as a number of other lucrative occupations at the extreme right end of the income distribution (Kerr et al., 2017). All of these determinants matter.

There are other known determinants of income disparities and, again, they are essential for understanding both individual and group differences. For example, full-time workers are often isolated for analytic purposes when group differences are examined statistically. However, full-time workers vary tremendously in their work experience and the numbers of hours they work and are willing to work within all professions. Just as the top 1% of ability contains one third of the ability range, one third of the range of how much time people devote to their careers is found among full-time workers (Lubinski & Benbow, 2021; Murray, 2003). In certain instances, taking hours devoted to work into account can be clarifying for understanding individual and group differences. For example, it has been found that professional-managerial men are 2.7 times as likely as equivalent women to work 50 hours or more per week (J. C. Williams & Boushey, 2010). It has also been documented that, due to differences in value placed on family-time, women find that they would actually need to be compensated more than men for each hour worked in order to maintain an equivalent level of happiness (Brockmann et al., 2018).

Awareness of important life preferences such as these helps one interpret and put into context phenomena like the gender wage gap.

To illustrate this point further, consider a 20-year longitudinal study conducted by Benbow et al. (2000). This study followed 1,995 mathematically precocious youth from ages 12 to 33. The authors regressed income on gender, while controlling for hours worked per week as well as the occupations participants worked in. Analyses were conducted within nine occupational categories: medical doctors, postsecondary teachers, engineers, lawyers, mathematicians and computer scientists, natural and physical scientists, executives and administrators, one "other-high prestige" category, and one "other-low prestige" category. The authors found no significant main effects of gender or any interactions with it in the prediction of income. Thus, by controlling for occupational category and hours worked per week, the authors were able to show effects of gender on income attained at early career was not significantly different from zero.

Many contemporary studies published in prestigious outlets neglect the extent to which these preexisting determinants can influence important outcomes. A recent study in the *Proceedings of the National Academy of Sciences* (Sterling et al., 2020), for example, discussed the gender wage gap as it relates to gender differences in confidence without ever mentioning other factors (e.g., gender differences in ability and interest patterns, work structure preferences). Factors such as the extent to which women and men choose different educational and occupational tracks, choose to enter less-than-glamorous work environments, work more hours (plus take on inflexible or unpredictable work schedules), and the extent to which *information decay* (McDowell, 1982) requires a continuous updating of complex skills all matter for understanding differential outcomes in compensation. All of these considerations make some

careers more lucrative than others (e.g., working as an on-call physician, working as a high-power cybersecurity specialist, working two-week-long shifts fracking for gas in northcentral Pennsylvania or on oil rigs in the Gulf of Mexico).

To illustrate another subtle way in which this literature can be interpretively difficult, consider the range of work-history experience routinely found in all occupations among full-time employees. Preliminary analyses conducted on four cohorts of the Study of Mathematically Precocious Youth (SMPY; detailed in the Methods section) reveal differences in work histories of current full-time workers. Figure 1 shows these results for three cohorts of intellectually precocious youth and one cohort of elite STEM doctoral students identified in 1992, when they were enrolled in the top 15 STEM doctoral training programs in the U.S. (all of whom are now around age 50); all four cohorts had exceptional ability levels. Across the fifteen-year interval leading up to age-50, women in the sample generally reported slightly fewer hours worked per week, relative to men. These differences, although not exceedingly large, compound over a career and lead to large differences in compensation over time. The same is true for other determinants such as willingness to travel extensively for work (Ferriman et al., 2009; Lubinski et al., 2014)

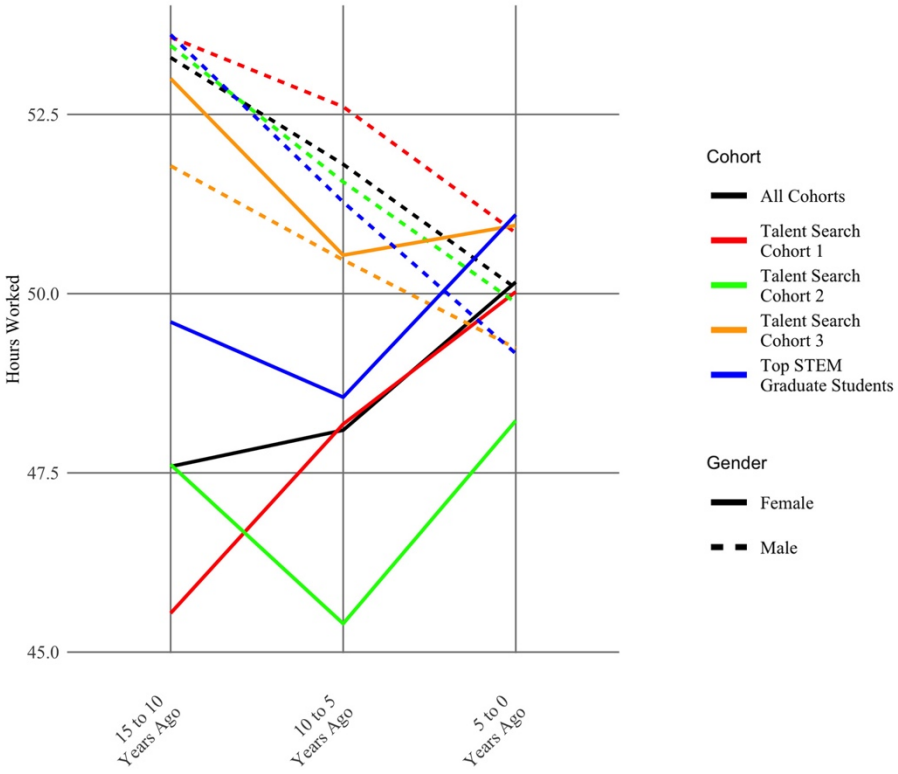

Figure 1. Hours Worked Retrospectively by Gender and Cohort.

The three talent search cohorts (Cohorts 1, 2, and 3) consist of age-50 data from participants from within the top 1%, 0.5%, and 0.01% of ability, respectively. The top STEM doctoral students were identified as first- or second-year doctoral students in 1992 enrolled in one of the top 15 STEM doctoral training programs in the United States. Among the talent search cohorts, 88% of the men and 66% of the women were working full time; among the top STEM doctoral students, 94% of the men and 76% of the women were working full time. Further details are found in the methods section.

A comprehensive overview of each of these determinants of differences in pay for the sexes is beyond the scope of this dissertation. Nevertheless, acknowledging and taking these determinants into account is required to advance knowledge on this topic, and yet their influence is frequently neglected (Ceci et al., 2014, 2021; Lubinski, 2010, 2016; Stewart-Williams & Halsey, 2021). Many of the studies and findings in the articles cited herein are the culmination of decades of scientific effort into investigating individual differences that are directly relevant to the seventh-ranked question in *Nature's* top ten list of pressing questions facing social scientists.

To attempt to model gender disparities without factoring in these robust antecedent determinants results in underdetermined models (Lubinski, 2010; Sterling et al., 2020).

The research conducted on the nuances and dimensions of the gender wage gap, as discussed above, helps explain disparate salaries between women and men. Insight is gained through findings that take into account that the gender wage gap has decreased over time as women earned advanced degrees and credentials commensurate with those of men; that those jobs with the most flexibility show the smallest gender wage gap; and that the adjusted gender wage gap is appreciably reduced when determinants such as hours worked, educational choices, and occupational choices are taken into account.

Yet, two other important factors play a role in understanding the unaccounted variance in the male-female income disparity that remains: discrimination in the workplace and the gender-differentiating effects of parenthood.

Discrimination

Despite efforts to combat a longstanding history of gender discrimination in the United States workforce (see The Equal Pay Act of 1963), it is commonly believed that the extent to which women are discriminated against in the workplace appreciably contributes to the male-female income disparity. Determining the degree of discrimination and ways to identify and eliminate it is therefore a contemporary focus.

Some studies point to discrimination as hiding within the unexplained variance between gender salaries after controlling for potential confounds (Budig & England, 2001; Corbett & Hill, 2012); for example, the 5% wage gap between women and men in the adjusted wage gap statistic could be the result of discrimination (as well as other unmeasured confounds, of course).

Cook et al.'s (2018) recent study of the gender wage gap of Uber drivers showed that, even in an environment in which gender discrimination *by the company* was unlikely due to blinding and unnegotiable pay, a 7% wage gap existed. Upon analysis, the study showed that the gap was explainable by differences in experience, driving speed, and taking riskier and more rewarding neighborhoods on the job (which, for good reason, women were less willing to take). Importantly, in this instance, *discrimination* did not explain the wage gap, but other factors do.

Meanwhile, compelling evidence has been put forth for differences in career success as a function of supply side characteristics of employees (both facilitating and limiting) as compared to demand side characteristics of occupational environments. Examples of supply side characteristics include employees' desire to seek positions of power and willingness to compete in the application pool for such positions, whereas demand side characteristics include discrimination in the workplace. W. M. Williams and Ceci (2015) found evidence of a 2:1 preference for equally-qualified female applicants over males for faculty positions among members in hiring committees and suggested that the largest hurdle to women in academia is actually *applying* to these positions in the first place (Ceci et al., 2014; National Research Council, 2010; Wolfinger et al., 2008). They did not find the same preference for less qualified women over more qualified men, however (Ceci & W. M. Williams, 2015). Gino and her colleagues (2015) showed that women have lower appraisals of how favorable positions of power would be for themselves—another supply side characteristic (due to the time demands and unpredictability of work schedules). Importantly, this was not due to thoughts or feelings that they could not do the job; they indeed felt that they could but would rather not over the long haul. Hence, Gino et al. (2015, p. 12354) title their *PNAS* article, "Compared to men, women view professional advancement as equally attainable, but less desirable". Concurring with this

appraisal and in commenting on the gender differences found among a series of longitudinal studies based in hundreds of intellectually brilliant women and men—which found that intellectually prodigious women were more likely to value a somewhat more balanced life across community, family, friends, and work—Murray (2020, p. 78) remarked to readers: "If you try to argue that these women were duped into accepting traditional female roles, you run into a problem: Chances are that the women who made those judgments are a lot smarter than you are."

Historically, of course, there have been many examples of direct evidence for gender discrimination. One well-known example of direct evidence for discrimination comes from Neumark's (1996) study which uncovered evidence of discrimination when gender-identifying first names were attached to otherwise identical resumes and sent out to employers at restaurants; higher-paying restaurants preferred men over the women. Similar experimental results were demonstrated by Correll et al. (2007). Further direct evidence of discrimination comes from Goldin and Rouse's (2000) study showing that jurors of orchestras were more likely to hire women when musicians' genders were blinded in auditions than when they were able to see the gender of the musician.

Taken holistically, the body of contemporary evidence seems to suggest that discrimination between the genders in hiring and remuneration does exist, but that the systematic effects of gender discrimination on pay is small. The situation has improved over time, with many of the most flagrant examples of discrimination becoming dated (Correll et al., 2007; Goldin & Rouse, 2000; Neumark et al., 1996) and subsequently corrected to an appreciable degree. Because the gender wage gap is largely closed by including these important determinants, the systematic effect of discrimination, though not trivial, must be smaller than most would suspect.

Overall, much progress has been made over the past two decades, although discrimination undoubtedly still exists in some disciplines and domains. In general, the literature suggests a view not dissimilar from Claude Canizares's remarks following his leadership role as Co-Chair of the U.S. National Academies Report (2010), "Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty." In an interview conducted by *Science*, wherein a thumbnail sketch of the empirical findings of this report were outlined, Canizares remarked, "While women can take some encouragement from the fact that there is no evidence of large-scale bias at these key transition points [across STEM careers], the reasons for their continued underrepresentation need to be examined more closely" (Mervis, 2009, p. 1250). Furthermore, Canizares went on to suggest that in order to better understand this phenomenon, federal agencies and universities should collect longitudinal data on the career paths of women—and suggested that we start with our own graduate students. Following Canizares's suggestion, this dissertation will provide just this type of longitudinal data on the career and life outcomes of three cohorts of intellectually talented women and men totaling 1,952 participants tracked for 35 years. It also will include a fourth cohort situated expressly to address Canizares's focal concern about STEM careers: 522 STEM doctoral students who attended the top 15 STEM doctoral training programs in the U.S. in 1992 (48% women) and were subsequently tracked for 25 years. Before proceeding to the particulars of the multidecade, four-cohort longitudinal study that I am examining, one final purported determinant of the income gap between the genders requires review. The empirical evidence suggests that it is an especially potent one.

Children and Parenthood

The final determinant of the gender wage gap to be reviewed is the presence of children and parenthood. It has been observed (and even reported in popular press in outlets such as *The New York Times*) that with each additional child in the household, mothers' salaries decrease; meanwhile, fathers' salaries remain constant or even *increase* with each additional child (Miller, 2014). The former phenomenon has been called the "Motherhood Penalty", and the latter has been called the "Fatherhood Premium" or "Fatherhood Bonus" (but see pages 22-23 for a discussion of the value-laden nature of this terminology and the discourse that typically accompanies it). Similar effects have been found for mothers' and fathers' work productivity (W. M. Williams & Ceci, 2012). Some have gone as far as saying that the gender wage gap that remains today can largely be explained as one of many gender differentiating effects of parenthood (Kleven et al., 2019). While the wage reductions accompanying motherhood have long-since been recognized (even taken for granted), there has been a plethora of studies over the last 25 years in disciplines ranging from economics (Azmat & Ferrer, 2017; Bertrand et al., 2010; Goldin, 2014; Sasser, 2005; Schulze, 2015) to sociology (Budig & England, 2001; Noonan et al., 2005; Waldfogel, 1997) offering reasons why parenthood is associated with a gender wage gap. These reasons do not only inform our understanding of the gender wage gap. They also inform (and have been informed by) theorizing in evolutionary psychology (Browne, 2002, 2005; Buss, 2019; Geary, 2021) and economics (Hakim, 2017) about anticipated gender differences over the lifespan. Ultimately, they also give rise to interesting hypotheses regarding which *types* of individuals and careers are likely to be affected by the most pronounced pay differentials as a function of parenthood.

In 2001, Michelle Budig and Paula England published a study in the *American Sociological Review* based on the 1982-1993 National Longitudinal Survey of Youth; they concluded that mothers experience a wage penalty of 7% per child. Controlling for job experience reduced the penalty somewhat, but characteristics of the job did not explain much of the penalty at all. Not long before, Waldfogel (1997) determined that the Motherhood Penalty could be partially, but not entirely, attributed to differences in labor market experience. These findings suggested that some, but not all, of the gender wage gap could be attributed to differences between males and females in labor market experience; extrapolating suggests that women with the highest earning potentials should be affected by the highest percentage of decrements in pay.

Subsequently, empirical studies began to show especially pronounced decrements in salaries for mothers in especially demanding, high-powered work environments. Studies of physicians (Sasser, 2005), lawyers (Azmat & Ferrer, 2017; Noonan et al., 2005), MBAs (Bertrand et al., 2010), and PHDs (Schulze, 2015) all pointed to motherhood as engendering reductions to a woman's career prospects in the most competitive and least flexible professions. These results again intimated that certain types of women (particularly those who are most capable of qualifying for and securing the largest salaries) are those who are most likely to be affected by the Motherhood Penalty.

Recent work by Yu and Kuo (2017) and Weeden and colleagues (2016) discussed again that the Motherhood Penalty is particularly large in the least flexible, most demanding occupations, re-emphasizing the need to examine the extent of the Motherhood Penalty in highly skilled women. Goldin (2014, p. 1092) has said that, "The gender gap in pay would be considerably reduced and might vanish if firms did not have an incentive to disproportionately

reward individuals who worked long hours and who worked particular hours". But remarks like this are somewhat perplexing, because, while unquestionably true, some positions simply require long and erratic hours and impose inconvenient demands due to unexpected exigencies that characterize occupational roles that entrust people with vast amounts of economic and human resources. This is expressly why inordinately demanding positions are highly compensated. Across leadership positions in industry, law enforcement, law, medicine, the military, and university administration (among others), it is unlikely that these demanding challenges associated with highly compensated positions will go away. Some occupational roles not only need the best and the brightest, but also the atypically committed who possess a willingness to perform on short notice.

Contemporary meta-analytic findings have emerged to summarize this literature. Cukrowska-Torzewska and Matysiak's (2020) meta-analysis estimated an overall motherhood wage gap of between 3.6% and 3.8%. The authors were able to attribute the gap predominately to losses in mothers' human capital (e.g., expertise development and sharpening and work experience) and choices in occupation following birth of the first child.

The collection of studies conducted thus far on the gender wage gap indicates that, with credentialing having reached parity and the amount of discrimination estimated to be relatively small, particular emphasis should be given to the differential effects of parenthood on mothers and fathers. Moreover, the body of research just reviewed hints at the importance of studying high-potential women and men with inordinate promise for professional success. Having given an overview of the dimensions underlying the gender wage gap as well as some of the commonly theorized causes for its existence, I now focus the remainder of this dissertation on the broad-spectrum collateral consequences of parenthood and the extent to which they operate among

society's most intellectually able individuals and the subsets of those who have received truly outstanding educational and career opportunities.

An Emphasis on High-Potential Populations

In recent years, an interest in the extent to which the Motherhood Penalty and Fatherhood Premium operate in high-potential populations has intensified. High-potential individuals are in demand in today's conceptual society (Friedman, 2005; Hunt, 1995, 2011; Zakaria, 2008); in the words of former Labor Secretary, Robert Reich (1991), these "symbol analysts" are in demand for lucrative employment in modern conceptual economies. Intellectually talented individuals often find themselves entrusted with positions of responsibility and power; however, as noted earlier, these positions frequently require demanding, inflexible schedules as well as a constant updating of the incumbent's skills and competencies (Harari, 2018; Hunt, 1995; McDowell, 1982). So, high-potential populations have the advantage of many choices in the world of work given the demand for their talents. They also in general have many more choices to structure their lives and the wherewithal to develop in accordance with their preferences (Ferriman et al., 2009; Lubinski et al., 2014; Murray, 2020; Susan Pinker, 2008; Rhoads, 2004). Given their human capital, which affords them many choices, and their demand in modern economies, high-potential population are particularly interesting for examining how tradeoffs are navigated between economic gain and lifestyle preferences and priorities.

Society is just getting to a place where we can begin to study the choices and tradeoffs high-potential women make over their life course because historically their options in the world of work were so limited. Before second-wave feminism, intellectually talented women could be

executive assistants, nurses, and teachers, but not CEOs, doctors, and professors. While there is certainly room for further progress in terms of equal opportunity, contemporary times are markedly different than in the past. While all complex societies are constantly evolving, sufficient time for societal readjustments and varying lifestyles to be measured against the backdrop of an expanded range of progress toward equal opportunity have become successively more meaningful.

At the same time, G. S. Becker's (1985, 1991) Human Capital Theory predicts that parenthood diminishes wages of mothers because it prevents them from developing their employment potential; due to time constraints, the high-level skills relevant to success in high-power professions cannot be fully updated and optimally actualized. This theory also posits that the parent with more responsibility for raising the family (typically the mother; Sayer et al., 2004), will expend more energy at home and thus will have fewer resources to devote to career development, ultimately resulting in lower pay (Gough & Noonan, 2013). Childbearing and child rearing presumably detract from developing the skill sets necessary to succeed in cutting-edge careers, and this is especially critical during early career development (Taniguchi, 1999).

Drawing on the conceptual nature of the economy today and G. S. Becker's (1985, 1991) Human Capital Theory, investigators have wondered whether the Motherhood Penalty within high-potential samples is especially pronounced, relative to the general population. Specifically, some have hypothesized that high-potential women who disproportionately commit to raising children within the household, relative to their partners, are affected by a substantial decrement in occupational outcomes inasmuch as commitments outside of work make it more difficult to keep up with the rapid pace of demanding careers. I turn now to review the evidence.

Researchers have questioned whether women who have the potential to earn the most also would be the ones who experience the largest wage penalties following parenthood. In 2010, Michelle Budig and Melissa Hodges began an early investigation of the Motherhood Penalty across low-wage, middle-wage, and high-wage workers. Using NLSY data between 1979 and 2004, the authors concluded that the Motherhood Penalty was largest on low-wage women (although it existed for women at all levels). The methodology of this study, however, was criticized by Killewald and Bearak (2014) who concluded that, using the same data set, it was actually women at the *middle* of the wage distribution who suffered the largest penalties for having children. Finally, in a follow-up including co-authorship of a number of authors on these two conflicting publications, England and colleagues (2016) using NLSY data from 1979 to 2010, determined that it was actually the most highly skilled and highly paid women who suffered the greatest financial penalties subsequent to having children. Meanwhile, to complicate matters further, Glauber (2018) determined in the Current Population Survey that, by the early 2010s, high-earning women paid a lower Motherhood Penalty compared to lower earning women (or even none at all).

Researchers have also explored differences in wage gaps for women with different educational levels. In an early example, Anderson and colleagues (2002) determined, using the National Longitudinal Survey of Labor Market Experience of Young Women, that white women with more years of education paid a larger Motherhood Penalty than did those with less education; this was largely explainable by time off of work (recall Figure 1 above). Wilde and colleagues (2010) produced similar results that those women who had more education—and who tended to have children later in life—had the greatest Motherhood Penalty. On the other hand, Todd (2001), using the Luxembourg Income Study, and Anderson and colleagues (2003), using

data from the National Longitudinal Survey of Young Women, have found that women with medium-educational attainment (high school graduates) actually suffered the greatest Motherhood Penalty.

To say that an overview of this literature—focusing on variation of the Motherhood Penalty across low- and high-potential populations—is confusing would be an understatement. With the same investigators coming to different conclusions using similar data sets (Anderson et al., 2002, 2003; Budig & Hodges, 2010; England et al., 2016; Killewald & Bearak, 2014), one would be hard-pressed to truly understand the extent of the gender wage gap on high-potential women.

This dissertation investigates (among other things) whether those women who are most at promise experience a substantial wage decrement for motherhood, relative to such men for fatherhood. Importantly, I will also go beyond documenting the extent to which this occurs. Namely, in addition to examining parents' compensation as a function of having children and the number of children they have, I will simultaneously examine other valued aspects of their lives which unfold in tandem. To my knowledge, no other study has examined male-female income divergences as a function of parenthood or number of children in the context of other highly valued aspects of life. Therefore, along with the dynamics of income changes, I will concurrently examine in what ways, if any, changes occur following parenthood in levels of psychological well-being, relationship satisfaction, and satisfaction with life more generally.

CHAPTER II

CURRENT INVESTIGATION

In much of the literature on the Gender Wage Gap and Motherhood Penalty, there is an intermixing of description and prescription. Nearly exclusively, wage differences between men and women are the sole outcome variable and, when disparities are observed, Discussion sections typically contain ways to attenuate the observed disparities. One reason the literature review provided above is helpful is because it clarifies known determinants of gender differences in income, and the personal attributes and work requirements that need to be factored into interventions and opportunities to offset these differences, if they are deemed appropriate. In some literature, it is unclear whether authors are recommending changes to individual orientations or changes to the work environment, and it is frequently unclear how some work environments could be changed to accommodate individuals who prefer not to pursue highly compensated occupational roles that encroach upon one's lifestyle more than typical occupations (Cheryan et al., 2017; Cheryan & Markus, 2020; El-Hout et al., 2021).

Indeed, even some of the terms used to describe the phenomenon are valued laden with "bonus" and "penalty" routinely employed when examining a unidimensional outcome such as monetary compensation at work. The current investigation will try to eschew this rhetoric by characterizing compensation and status changes at work in the context of other highly valued personal and psychological outcomes. The objective is to test the hypothesis that income divergences occur for high-potential men and women as the literature suggests, while

simultaneously examining other dynamic changes in the lives of highly able participants as they are in the process of building their families. The goal is to more fully understand the extent to which income disparities, if they should occur, covary with other valued aspect of life and in what direction. There is precedent for examining multiple valued outcomes over the developmental trajectory of high-potential individuals (because demonstrable life events cascade multiple life changes), and assessing potential changes multidimensionally can be psychological clarifying.

For example, in a 40-year longitudinal study of 3,467 mathematically gifted adolescents, Park and colleagues (2013) were interested in the extent to which appropriate developmental placement ("acceleration") was ultimately related to occupational and creative contributions in STEM (e.g., advanced degrees, occupations, patents, refereed STEM publications). The investigators did indeed find that accelerative learning experiences were significantly related to these important outcomes decades later, but there was one important exception: this relationship only appeared to be operative for males. This was a perplexing finding, because the girls did just as well as the boys in these experiences and, interestingly, if anything, at the time they appeared to value them more, which is a commonly observed gender difference among intellectually talented youth (Benbow & Stanley, 1996; Lubinski & Benbow, 2021). Upon further analysis, however, a broadening of the outcome space clarified what was happening. The women were much more likely than the men to specialize in medicine and law relative to inorganic STEM areas. Therefore, the criteria selected to evaluate the educational efficacy of acceleration for ultimate educational, occupational and creative outcomes were largely irrelevant to their individuality and chosen life paths. These women were not under-achieving. Rather, they were achieving in other areas and disciplines.

So too, because highly valued outcomes are seldom examined when studying dynamic gender differences in income as a function of parenthood and number of children, this investigation will do that here. Moreover, while theorists associated with conceptual frameworks related to the Motherhood Penalty and Fatherhood Bonus have been fairly taciturn with respect to important aspects of life beyond income, Hakim's (2017) Preference Theory and evolutionary psychologists (Browne, 2002, 2005; Buss, 2019; Geary, 2021) have speculated that gender differences in income and stature in the world of work following parenthood should not necessarily be expected to display the same pattern in other aspects of life. They hypothesize that as children come along there is a tendency for men and women to diverge some in work/family priorities. In general, men tend to focus more on their work (and tangible support for the family, through resources acquisition) while women in general tend to focus more on hands-on home and family (emotional support in times of need). Therefore, specifically, I will investigate the following three sets of hypotheses inspired by this literature:

H_{A1}) High-potential mothers will experience a substantive and significant decrease in wages as a function of the number of children they have; high-potential fathers will experience no decrease (or will even experience an increase). That is, I hypothesize an interaction effect[1].

H_{A2}) High-potential mothers will experience a decrease in the probability of being occupational leaders as a function of the number of children they have; high-potential fathers will experience no decrease (or will even experience an increase). That is, I hypothesize an interaction effect.

[1]Throughout this dissertation, I do not use the term "effect" to imply causal effect. The data collected here generally do not speak to causality in these hypotheses; the reader could easily come up with alternative causal mechanisms for each hypothesis.

H_{B1}) High-potential mothers at age 50 will experience working fewer hours as a function of the number of children they have; high-potential fathers will experience no decrease (or will even experience an increase). That is, I hypothesize an interaction effect.

H_{B2}) High-potential mothers at age 50 will experience spending more hours with family and the home as a function of the number of children they have; high-potential fathers will experience no increase (or will even experience a decrease). That is, I hypothesize an interaction effect.

H_{B3}) High-potential mothers at age 50 will report a willingness to work fewer hours, if given their ideal job, as a function of the number of children they have; high-potential fathers will experience no decrease (or will even experience an increase). That is, I hypothesize an interaction effect.

H_{C1}) Both high-potential mothers and fathers will experience higher psychological well-being as a function the number of children they have. That is, I hypothesize no interaction effect.

H_{C2}) Among those in relationships, both high-potential mothers and fathers will experience greater relationship satisfaction as a function of the number of children they have. That is, I hypothesize no interaction effect.

H_{C3}) Both high-potential mothers and fathers will experience higher life satisfaction as a function of the number of children they have. That is, I hypothesize no interaction effect.

Study 1

Methods

Participants

Participants in this study were drawn from SMPY (Lubinski & Benbow, 2006). SMPY was founded in 1971 as a planned 5-year study by Dr. Julian C. Stanley at The Johns Hopkins University under a grant from the Spencer Foundation (Stanley, 1977). The study has since been adapted into a 50-year longitudinal study and is currently co-directed by Drs. Camilla P. Benbow and David Lubinski at Peabody College of Vanderbilt University. SMPY has identified and longitudinally tracked over 5,000 intellectually talented youth from five cohorts over multiple decades (Clynes, 2016; Lubinski, 2016; Lubinski & Benbow, 2006, 2021).

Cohort 1 is comprised of individuals who were identified between 1972 and 1974 as gifted 13-year-olds. These participants scored in the top 1% of cognitive ability according to above-level standardized tests (SAT-M $\geq$ 390 or SAT-Verbal $\geq$ 370). This sample came primarily from Maryland and consisted of 447 females and 687 males (n = 1,134); the sample was 95% white or Caucasian, 0.8% black or African American, 0.5% Hispanic, 1.7% Asian, Asian-American, or Pacific Islander, and 1.9% Other.

Cohort 2 is comprised of individuals who were identified between 1976 and 1979 as highly gifted 13-year-olds. These participants scored in the top 0.5% of cognitive ability according to above-level standardized tests (SAT-M $\geq$ 500 or SAT-Verbal $\geq$ 430). This sample came from throughout the mid-Atlantic states and consisted of 160 females and 327 males (n = 487); the sample was 88.7% white or Caucasian, 0.4% black or African American, 0.8% Hispanic, 7.0% Asian, Asian-American, or Pacific Islander, and 3.1% Other.

Cohort 3 is comprised of individuals who were identified between 1980 and 1983 as profoundly gifted 13-year-olds. These participants scored in the top 0.01% of cognitive ability according to above-level standardized tests (SAT-M ≥ 700 or SAT-Verbal ≥ 630). This sample came from throughout the United States and consisted of 71 females and 260 males (n = 331); the sample was 75.8% white or Caucasian, 0.6% black or African American, 0.3% Hispanic, 18.8% Asian, Asian-American, or Pacific Islander, and 4.5% Other.

Procedure

Participants in Cohorts 1, 2, and 3 were initially identified and surveyed at age 13. They were identified through performance on *above-level testing* in which 13-year-olds were given college entrance exams designed for 17-year-old college-bound students (Stanley, 1990; Warne, 2012). Follow-up surveys were completed at ages 18, 23, 33, and 50. All participants included in this study were surveyed at least at age 13 (identification survey) and age 50 (mid-career follow-up). Response rates to the mid-career survey were 88.5%, 88.6%, and 75.9% for Cohorts 1, 2, and 3, respectively.

The mid-career follow-up survey was conducted in two phases. Cohorts 1 and 2 completed the mid-career follow-up between 2012 and 2013; Cohort 3 completed the follow-up between 2017 and 2018. Among other things, the mid-career follow-up asked participants questions regarding their and their significant other's occupational status; their gender identification at the time of the survey; number of children; preferences for work structure; and five questionnaires related to psychological well-being. As an incentive to complete the survey, participants were rewarded with either a $20 Amazon gift card or the opportunity to donate the equivalent amount to summer residential programs for intellectually gifted youth from disadvantaged backgrounds (67% chose to donate).

Instrumentation

Age-50 Survey

Two separate (but related) indicators of occupational achievement at mid-career were assessed in this study in order to broadly map occupational success: self-reported annual income and occupational leadership. Annual income (for both participants and their significant others) was reported by the participant during the mid-career follow-up.

To better understand participants' work structure and develop a more complete understanding of how SMPY participants structure their lives, two questions were asked on the mid-career survey regarding hours worked and hours willing to work. The number of hours currently worked by participants was given by asking participants to describe how many hours per week they worked, on average, over the five-year interval leading up to the completion of the survey. The number of hours willing to work in their ideal job was written in by participants responding to the question, "If you were given the opportunity to work in your job of first choice, how many hours per week would you be willing to work at most?". Participants were also asked to indicate how much time they spend with their family (including relatives) as well as maintaining a home, and these were combined into a measure of time spent dedicated to the family.

In order to broadly capture the breadth of psychological well-being, participants were assessed on five different indicators of psychological well-being in the mid-career follow-up. Each of these indicators is well-known. Moreover, each indicator assesses at least one of the two conceptualizations of psychological well-being: the eudaimonic and hedonic perspectives (Ryff et al., 2004).

Core Self-Evaluations (12 items; Judge et al., 2003) assesses one's evaluation of oneself and one's abilities. This scale contains four personality dimensions (locus of control, neuroticism, generalized self-efficacy, and self-esteem). For example, "When I try, I generally succeed". Responses range from strongly disagree (1) to strongly agree (5). Alpha reliabilities: Cohort 1 = .86, Cohort 2 = .88, and Cohort 3 = .89.

Psychological Flourishing (8 items; Diener et al., 2010) measures self-perceived success in relationships, self-esteem, purpose, and optimism. This scale measures social-psychological prosperity and also complements other measures of well-being. For example, "I lead a purposeful and meaningful life". Responses range from strongly disagree (1) to strongly agree (7). Alpha reliabilities: Cohort 1 = .85, Cohort 2 = .89, and Cohort 3 = .86.

Positive Affect (5 items; Diener et al., 2010), from the Scale of Positive and Negative Experience, asks participants about their frequency of positive feelings. Participants were asked to rate feelings of Positive, Good, Pleasant, Contented, and Happy. Responses range from very rarely or never (1) to very often or always (5). Alpha reliabilities: Cohort 1 = .90, Cohort 2 = .90, and Cohort 3 = .90.

Negative Affect (reversed; 10 items; Goldberg, 1992) was drawn from the International Personality Item Pool Big-Five 50-item inventory. This scale asks participants to rate on a Likert-type scale the extent to which a series of statements accurately describes them. This results in scores on each of the Big Five personality dimensions, but only the composite reflecting Negative Affect was used in this study. For example, "I often feel blue". Responses range from very inaccurately (1) to very accurately (7). Alpha reliabilities: Cohort 1 = .88, Cohort 2 = .88, and Cohort 3 = .91.

Relationship Satisfaction. Participants who were in relationships were asked to rate their overall satisfaction with their relationship on a one-item, seven-point scale: 1 = very dissatisfied, 7 = very satisfied.

Life Satisfaction (5 items; Diener et al., 1985; Pavot & Diener, 1993) measures global judgements of satisfaction with one's life. The scale does not ask participants about certain aspects of lives with which they are satisfied (e.g., finances); instead, this scale allows participants to evaluate their lives holistically, giving differential weight to each aspect of their lives, as they see fit. For example, "In most ways my life is close to my ideal". Responses range from strongly disagree (1) to strongly agree (7). Alpha reliabilities: Cohort 1 = .90, Cohort 2 = .90, and Cohort 3 = .90.

Participants were also asked to indicate their gender (not sex) during the mid-career survey; their options were "male", "female", and "other" (in the latter case, they could write in a response). All participants in this study indicated either "male" or "female". Participants also indicated the number of children and the status of each child (i.e., biological child, adopted child, or stepchild). For this dissertation, analyses will be restricted to biological children due to small numbers of other types of children and in order to facilitate cleaner interpretations of results.

Figure 2 shows the proportion of participants, by gender, who had each number of biological children in for Study 1. Due to sparsity of data in cells beyond six biological children, instability of parameter estimates that arises from sparsity, and issues with extrapolating inferences into regions that are not well supported by the observed data, all categories equal to and above six children were simply collapsed into one category ("6").

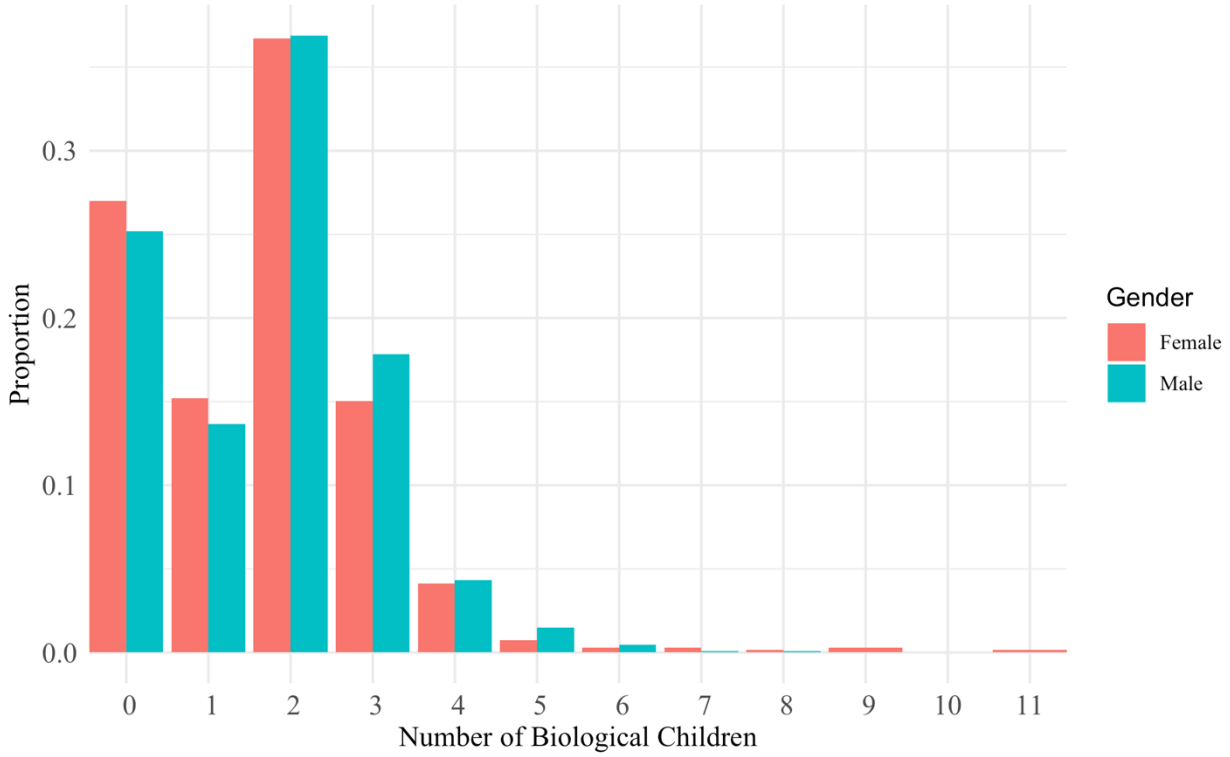

Figure 2. Distribution of Biological Children for Talent Search Participants.

L-Data (Life Record Data)

Occupational Leadership. Occupational leadership was assessed objectively. Participants were defined as occupational leaders if they had met the standard for what content experts considered *ultimate criteria* for occupational distinction (Simonton, 2014; Thorndike, 1949, pp. 121-124). In order to define ultimate criteria deemed truly impressive, we consulted the literature (Bernstein et al., 2019; K. O. McCabe et al., 2020; Murray, 2003; Simonton, 2014; Zuckerman, 1977) as well as distinguished consultants who were themselves leaders in their respective disciplines (an astrophysicist at Harvard, Vanderbilt's Dean of Law, Vanderbilt's Dean of Engineering, and Vanderbilt's government liaison). The idea was to identify a level of career accomplishment, at minimum, commensurate with securing tenure at a research-one university.

Participants were coded as occupational leaders if they had achieved one of the following criteria: tenure in a research-intensive university, an executive role in a Fortune 500 company, a

position of leadership in government (at or above GS-14), at least 20 patents, or over $2,750,000 in grant funding. Additionally, several participants were identified as occupational leaders based on other idiographic accomplishments considered worthy of meeting this standard of occupational distinction (e.g., winning prestigious awards such as a Pulitzer Prize) and others still were identified based on the recommendations of experts in the field with whom we consulted (for additional details, see Bernstein and colleagues, 2019 or K. O. McCabe and colleagues, 2020).

It's important to note that the information for this classification did not come from the mid-career surveys, but rather from web searches and Publish or Perish (Harzing, 2007) queries of participants conducted in 2016 and 2017. So, this outcome was assessed objectively as a part of their life-record (L) data (Cattell, 1957).

Analytic Approach

Analyses were conducted under the generalized linear modeling (GLM) framework (Hardin & Hilbe, 2018) with particular emphasis on multiple linear regression and logistic regression. Emphasis was placed on the interpretation of the interaction of participant gender and number of children in the prediction of the outcome of interest. This GLM interaction is interpreted in the general definition outlined by C. J. McCabe and colleagues (2021): the change in marginal effect of children for a counterfactual change from female to male. The (statistical) null hypothesis is taken to be that the interaction is zero in the population (i.e., number of biological children relates to the outcome of interest in exactly the same way for both men and women). Techniques for probing interactions and computing simple slopes were implemented according to Aiken and West (1991) as well as Preacher and colleagues (2006). Biological children are the focus of these analyses due to sparse sample size limitations in the non-

biological categories and to facilitate cleaner connections to germane economic theory (Hakim, 2017) and the evolutionary psychology literature (Browne, 2002; Buss, 2019; Geary, 2021).

As is typical with income data, there was extreme positive skew in our sample (i.e., a few individuals made much more money than others in the cohort). To deal with this positive skew and the non-normality of regression residuals that it entailed (which is a violation of a fundamental assumption of multiple linear regression), income was log-transformed and inference was conducted on the logarithmic scale.

In order to reduce the dimensionality of the five indicators of psychological well-being, and enhance the construct validity of assessing global psychological well-being with parsimony, parallel analysis (Horn, 1965) and principal component analysis were used. Parallel analysis allowed me to demonstrate that the five indicators of psychological well-being shared a broad underlying construct across each cohort, and principal component analysis allowed me to derive an ideal linear combination of predictors and relate this linear combination to the predictors of interest in the GLM framework. I used a similar procedure in Bernstein and colleagues (2021).

All data processing, statistical analyses, and data visualization was done using R. Packages of particular importance in this analysis include Revelle's (2021) psych package (for parallel analysis and principal component analysis), C. J. McCabe's (2021) modglm package for analysis of interactions in logistic regression, and Wickham and colleagues' (2019) tidyverse package for general data wrangling, management, and visualization.

Ancillary Analyses

Finally, in order to better understand the story of the ways in which highly able men and women structure their work and family, as well as the ways in which they feel about the decisions they make, a few ancillary analyses are conducted. These analyses investigate the

extent to which male and female participants decide to divide with their significant other the task of working full time in order to earn income and support their family as children come along. Relatedly, an ancillary analysis of the total amount of family income (participant income, significant other income, and income from inheritance and investments) is conducted to see whether, as a team, a division of labor between participants and their significant others affected the economic viability of the family unit as a whole as more children come along.

Results

Occupational Outcomes

Table 1 summarizes the regression results of log wages as a function of number of children and gender for the talent search participants. Results are given across the three talent search cohorts, and the final column shows the most accurate parameter estimates for this model with the cohorts combined. Figure 3 shows the relationship graphically. Across the three cohorts of intellectually prodigious participants, women saw strong decrements to log income as a function of having more children; the parameter estimate in the cohorts combined was strong, significant, and negative (simple slope = -0.482, $t(1844) = -5.43$, $p < .001$). Nevertheless, the interaction term was strong and positive ($b = 0.754$, $t(1844) = 6.872$, $p < .001$), resulting in substantial *increases* in log income for men as a function of number of children (simple slope = 0.273, $t(1844) = 4.21$, $p < .001$).

Table 1: Modeling Log Participant Income for Talent Search Participants

	Dependent variable:			
	Log Participant Income			
	Cohort 1	Cohort 2	Cohort 3	Talent Search
Number of Biological Children	-0.41***	-0.86***	-0.25	-0.48***
	(-0.62, -0.21)	(-1.24, -0.48)	(-0.86, 0.36)	(-0.66, -0.31)
Gender	0.03	-0.24	0.34	0.03
	(-0.54, 0.60)	(-1.14, 0.66)	(-0.93, 1.61)	(-0.42, 0.47)
Interaction	0.65***	1.19***	0.55	0.75***
	(0.38, 0.92)	(0.74, 1.64)	(-0.13, 1.24)	(0.54, 0.97)
Constant	10.92***	11.12***	10.60***	10.91***
	(10.48, 11.36)	(10.37, 11.87)	(9.50, 11.70)	(10.55, 11.26)
Observations	1,071	461	316	1,848
Log Likelihood	-2,640.04	-1,130.18	-797.02	-4,570.66

Note. The first three columns show consistency in parameter estimates across three separate cohorts of intellectually precocious youth; the last column gives the best parameter estimates for the model fit with the cohorts combined. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

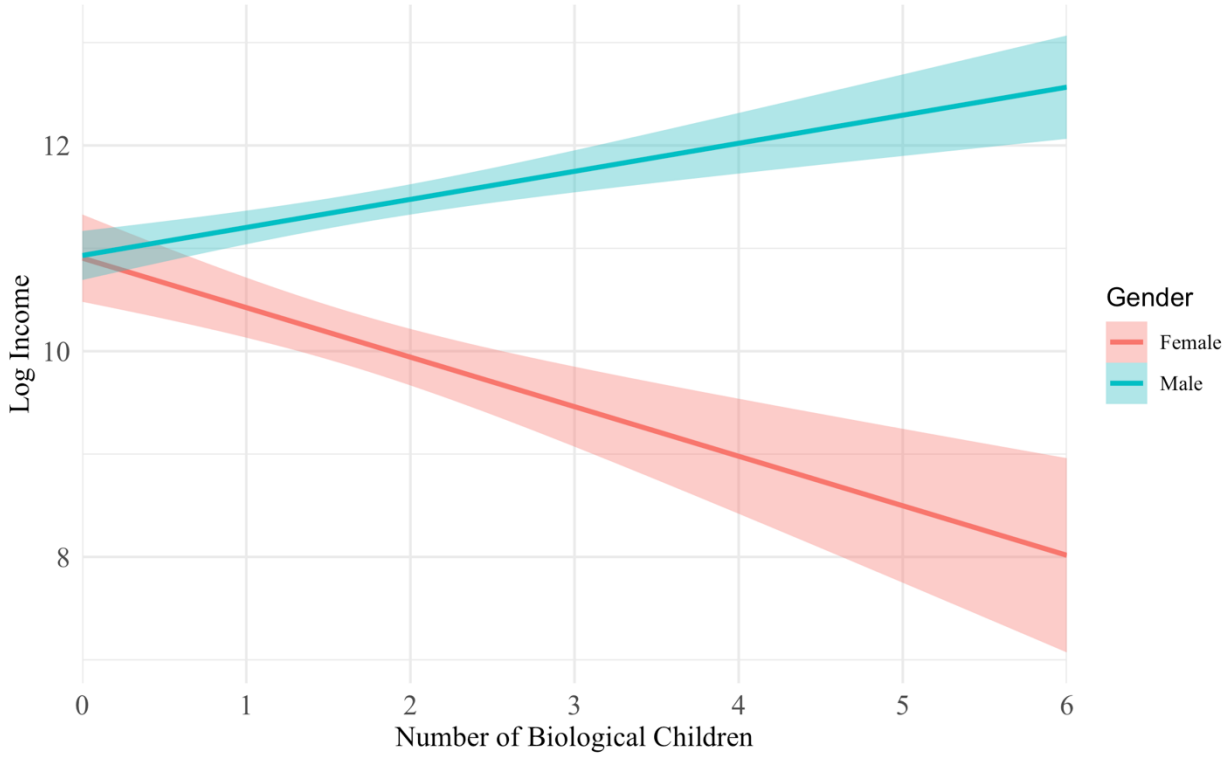

Figure 3. Study 1: Log Income as a Function of Number of Children.
Shaded bands give 95% confidence intervals about the regression line.

Figure 4 graphically represents the same model on the natural scale of interpretation: absolute income. On this scale, a one-unit increase in the number of children is associated with a 38% decrease in wages for women, whereas the same change is associated with a 31% increase in wages for men. The result is a conspicuous income divergence between women and men as a function of number of children.

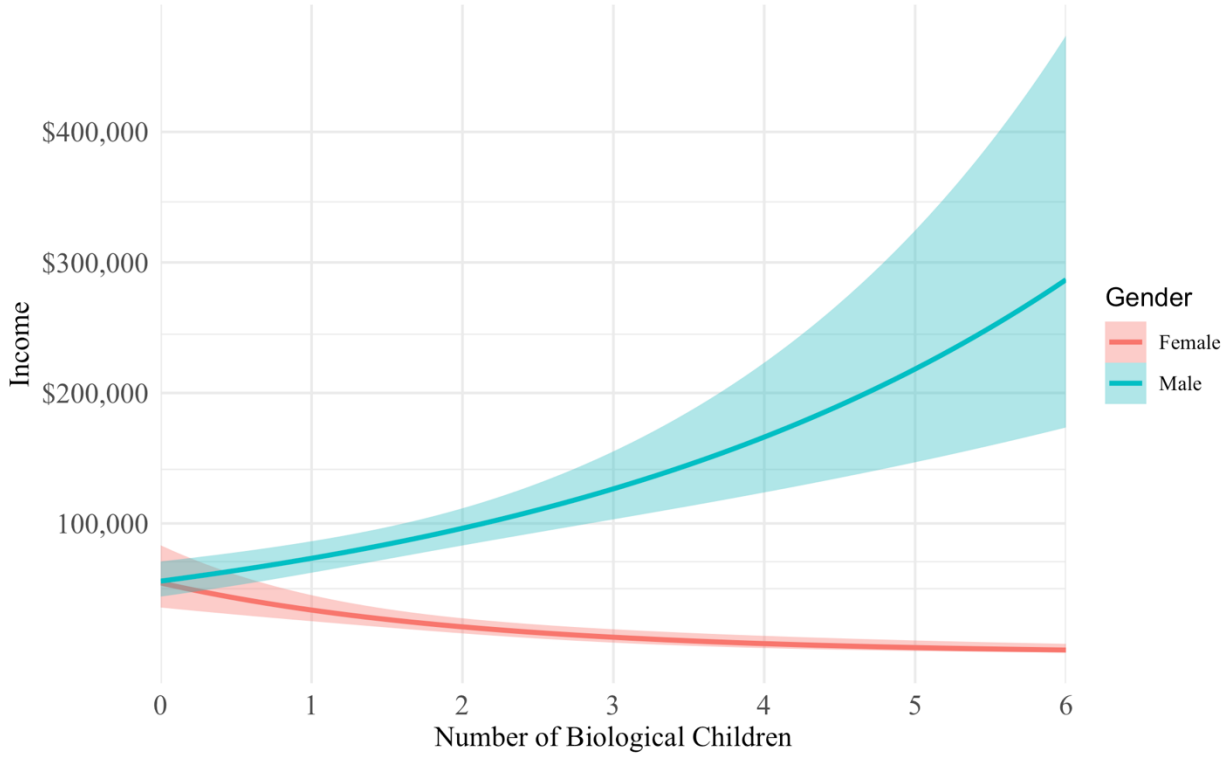

Figure 4. Study 1: Income as a Function of Number of Children.
Shaded bands give 95% confidence intervals about the exponentiated regression line.

Table 2 summarizes the results of the regression of log odds of leadership as a function of number of number of children by gender. Results are given across the three talent search cohorts, and the final column shows the most accurate parameter estimates for this model with the cohorts combined. Figure 5 shows the relationship graphically. Across the three cohorts of intellectually precocious youth, results were mixed. However, the model's parameter estimates based on the combined cohorts suggests a similar pattern observed in the analysis of (log) wages on the predictors: more children lowered the log-odds of leadership for women (simple slope = -0.048, $t(1948) = -0.497$, $p = .619$), and a positive interaction ($b = 0.210$, $t(1948) = 1.901$, $p = .0573$) reversed the trend for men (simple slope = 0.162, $t(1948) = 3.059$, $p = .002$).

Table 2: Modeling Log Odds of Leadership for Talent Search Participants

	Dependent variable:			
	Leadership			
	Cohort 1	Cohort 2	Cohort 3	Talent Search
Number of Biological Children	-0.08 (-0.34, 0.18)	0.14 (-0.25, 0.53)	-0.07 (-0.55, 0.42)	-0.05 (-0.24, 0.14)
Gender	0.51 (-0.11, 1.14)	0.60 (-0.32, 1.52)	-0.19 (-1.17, 0.79)	0.44* (-0.01, 0.88)
Interaction	0.27* (-0.03, 0.56)	-0.06 (-0.50, 0.38)	0.31 (-0.23, 0.84)	0.21* (-0.01, 0.43)
Constant	-2.27*** (-2.80, -1.75)	-2.12*** (-2.93, -1.32)	-1.15*** (-2.00, -0.29)	-2.02*** (-2.40, -1.64)
Observations	1,134	487	331	1,952
Log Likelihood	-462.01	-226.05	-190.90	-890.86

Note. The first three columns show consistency in parameter estimates across three separate cohorts of intellectually precocious youth; the last column gives the best parameter estimates for the model fit with the cohorts combined. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

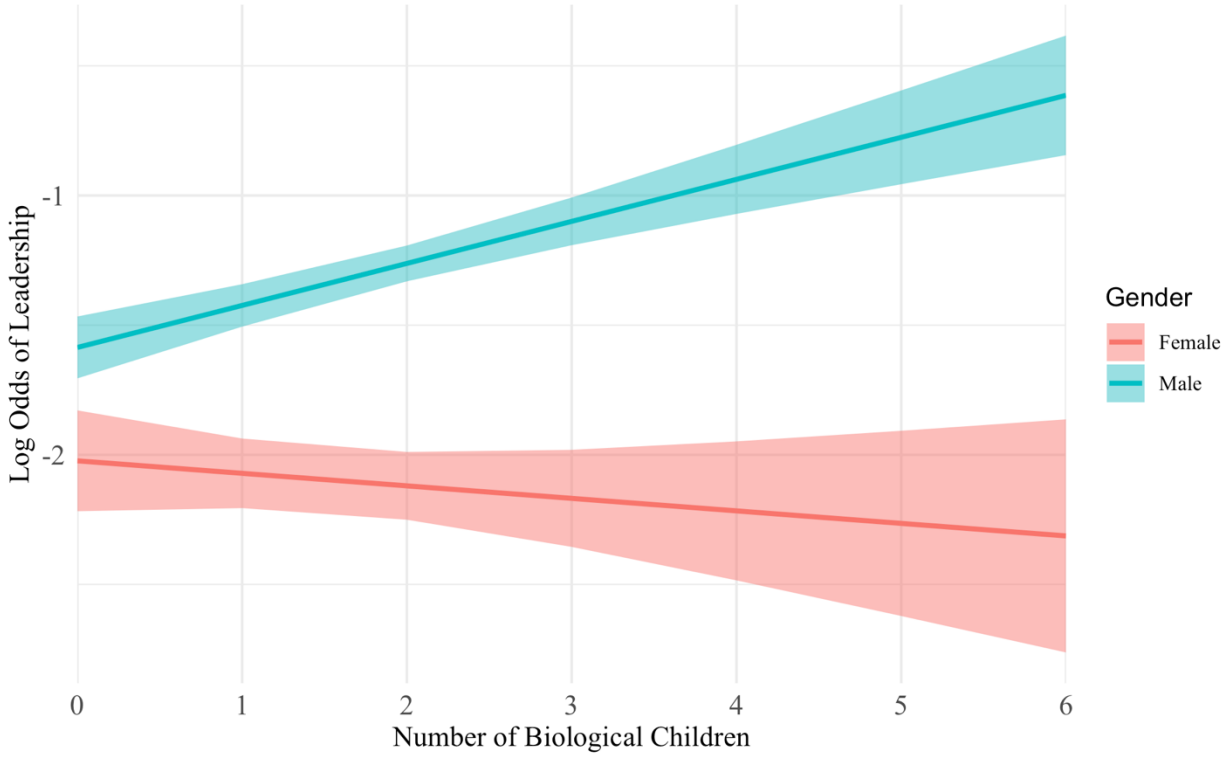

Figure 5. Study 1: Log Odds of Leadership as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

However, H_{A2} specifies that an interaction would be found on the *probability* of occupational leadership, not the log odds of leadership. Interpreting the product coefficient as evidence of an interaction is valid on the transformed scale of the outcome in a generalized linear model (e.g., log odds of leadership in a logistic regression analysis), but not on the natural scale (e.g., probability of leadership); most researchers simply fail to distinguish the two outcome scales when interpreting research findings (Ai & Norton, 2003; C. J. McCabe et al., 2021; Mize, 2019). C. J. McCabe and colleagues (2021) give a more general definition of an interaction as a change in a marginal effect of one variable as a function of change in another variable; they use partial derivatives and discrete differences to estimate these interactions while employing the delta method and bootstrapping to derive appropriate standard errors. Following their guidance and

implementing the tools in C. J. McCabe's (2021) modglm R package, I investigated interactions on the probability scale of occupational leadership.

Table 3 summarizes the gender-by-number-of-children interaction on the probability of leadership; Figure 6 presents the model graphically. The interaction is statistically significant at each value of number of children, with the interaction becoming larger as the number of children increases. Moreover, the effects are consistent with the interpretation on the log odds scale: more children lead to larger increases in probability of leadership for men relative to women.

Table 3: Interaction Estimates in Probability of Leadership for Talent Search Participants

# Biological Children	Interaction Estimate	Standard Error of Interaction Estimate	Statistically Significant?
0	.0278	.0122	Yes
1	.0301	.0125	Yes
2	.0325	.0130	Yes
3	.0348	.0138	Yes
4	.0371	.0146	Yes
5	.0391	.0154	Yes
6	.0409	.0159	Yes

Note. The interaction can be interpreted as the increased marginal effect of the number of biological children on probability of leadership for a change in gender from female to male.

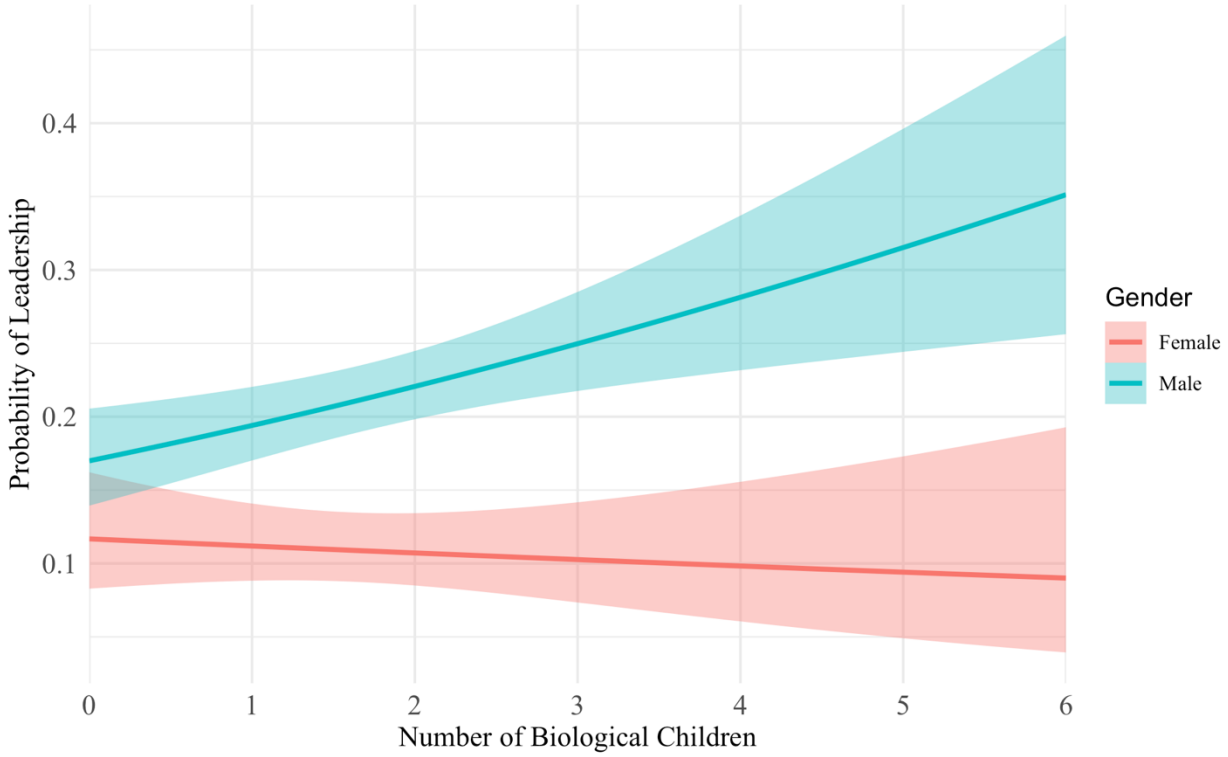

Figure 6. Study 1: Probability of Leadership as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Work-Life Structure

Table 4 summarizes the regression results of hours worked as a function of number of children by gender; Figure 7 shows the relationship graphically. Across the three cohorts, women's work hours decreased as a function of having more children; the cohort-combined parameter estimate was strong, significant, and negative (simple slope = -3.525, $t(1646) = -6.853$, $p < .001$). Moreover, the interaction term was strong and positive ($b = 4.004$, $t(1646) = 6.331$, $p < .001$), resulting in relative constancy in hours worked for men as a function of number of children (simple slope = 0.0.479, $t(1646) = 1.302$, $p = .193$).

Table 4: Modeling Hours Worked for Talent Search Participants

	Dependent variable:			
	Hours Working			
	Cohort 1	Cohort 2	Cohort 3	Talent Search
Number of Biological Children	-3.08***	-5.42***	-2.94	-3.52***
	(-4.31, -1.85)	(-7.39, -3.46)	(-6.58, 0.71)	(-4.53, -2.52)
Gender	2.29	-0.42	-3.63	0.50
	(-1.08, 5.66)	(-4.99, 4.15)	(-11.24, 3.98)	(-2.02, 3.06)
Interaction	2.88***	6.51***	4.42**	4.00***
	(1.29, 4.47)	(4.24, 8.79)	(0.30, 8.54)	(2.76, 5.24)
Constant	46.72***	46.60***	46.73***	46.56***
	(44.15, 49.30)	(42.77, 50.43)	(40.20, 53.26)	(44.52, 48.61)
Observations	958	425	267	1,650
Log Likelihood	-4,002.15	-1,707.80	-1,128.81	-6,855.97

Note. The first three columns show consistency in parameter estimates across three separate cohorts of intellectually precocious youth; the last column gives the best parameter estimates for the model fit with all cohorts combined. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

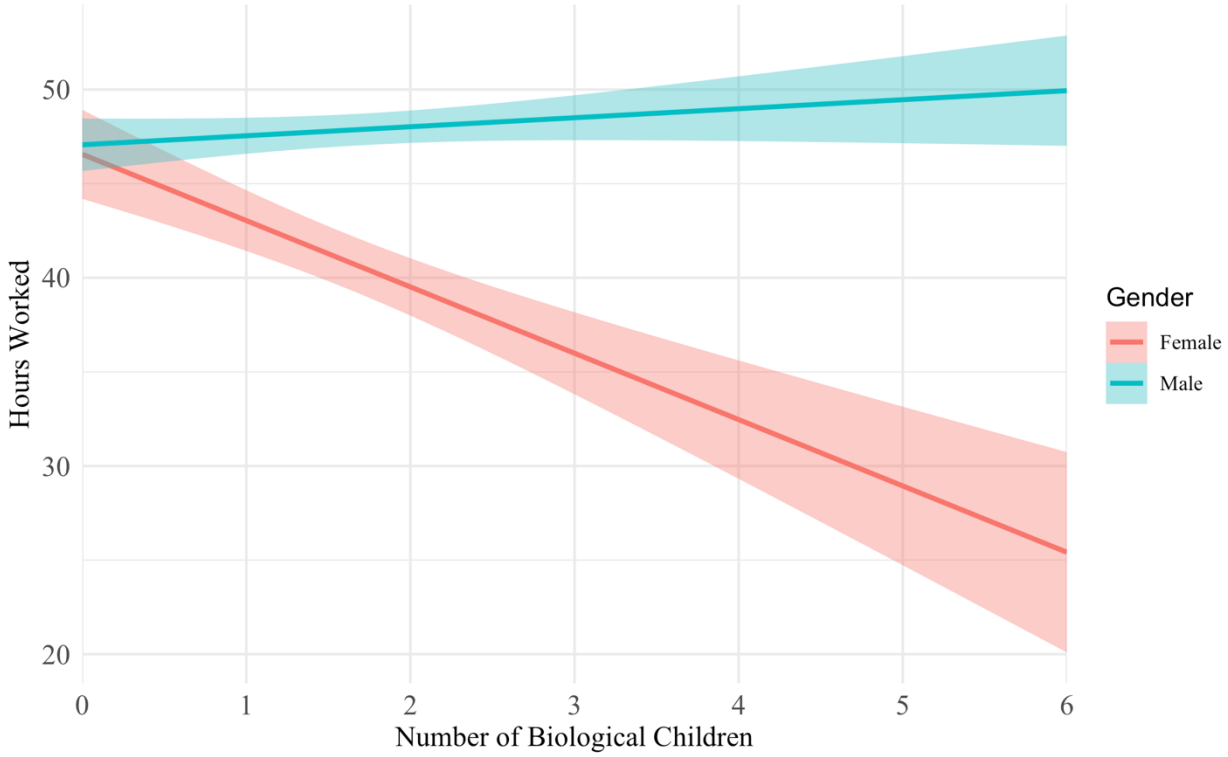

Figure 7. Study 1: Hours Worked as a Function of Number of Children.
Shaded bands give 95% confidence intervals about the regression line.

Table 5 summarizes the regression results of hours spent with family and home maintenance as a function of number of number of children by gender; Figure 8 shows the relationship graphically. Women saw strong increases to hours spent with family and home maintenance as a function of more children; the parameter estimate for the cohorts combined was strong, significant, and positive (simple slope = 7.985, $t(1594) = 8.873$, $p < .001$). Moreover, the interaction term was strong and negative ($b = -5.882$, $t(1594) = -5.310$, $p < .001$), resulting in less of an increase (though still an increase) in hours spent with family as a function of number of children for men (simple slope = 2.103, $t(1594) = 3.256$, $p = .001$).

Table 5. Modeling Hours with Family and Home for Talent Search Participants

	Dependent variable:			
	Hours with Family			
	Cohort 1	Cohort 2	Cohort 3	Talent Search
Number of Biological Children	5.12*** (2.95, 7.30)	17.54*** (13.72, 21.36)	7.11*** (2.26, 11.95)	7.99*** (6.22, 9.75)
Gender	-1.73 (-7.71, 4.24)	7.38 (-1.55, 16.31)	2.54 (-7.33, 12.40)	-0.30 (-4.80, 4.19)
Interaction	-3.29** (-6.10, -0.47)	-15.33*** (-19.77, -10.89)	-4.65* (-10.11, 0.80)	-5.88*** (-8.05, -3.71)
Constant	41.40*** (36.81, 45.98)	27.74*** (20.28, 35.19)	30.68*** (22.16, 39.19)	37.42*** (33.82, 41.02)
Observations	929	417	252	1,598
Log Likelihood	-4,391.43	-1,953.25	-1,120.52	-7,502.78

Note. The first three columns show consistency in parameter estimates across three separate cohorts of intellectually precocious youth; the last column gives the best parameter estimates for the model fit with all cohorts combined. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

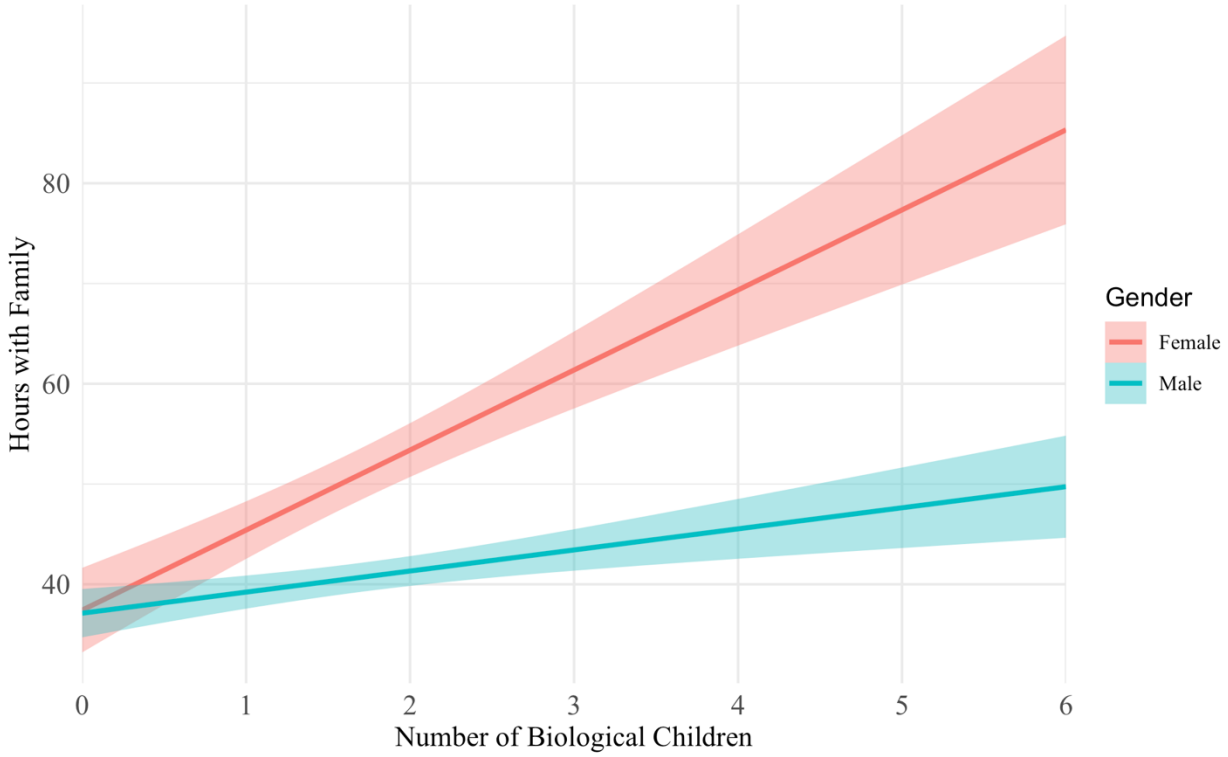

Figure 8. Study 1: Hours with Family as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Table 6 summarizes the regression results of hours willing to work in their ideal job as a function of number of number of children by gender; Figure 9 shows the relations graphically. Sharp decreases in hours willing to work was observed for women as a function of having more children; the parameter estimate in the cohorts combined was strong, significant, and negative (simple slope = -2.797, $t(1848) = -5.906$, $p < .001$). Moreover, the interaction term was strong and positive ($b = 2.478$, $t(1848) = 4.254$, $p < .001$), resulting in relative constancy in hours willing to work as a function of number of children for men (simple slope = -0.319, $t(1848) = 0.9404$, $p = .347$).

Table 6: Modeling Hours Willing to Work for Talent Search Participants

	Dependent variable:			
	Hours in Ideal Job			
	Cohort 1	Cohort 2	Cohort 3	Talent Search
Number of Biological Children	-2.39***	-3.54***	-3.85**	-2.80***
	(-3.49, -1.29)	(-5.38, -1.70)	(-7.46, -0.25)	(-3.73, -1.87)
Gender	5.73***	5.40**	1.40	5.26***
	(2.68, 8.78)	(1.07, 9.72)	(-6.14, 8.94)	(2.88, 7.64)
Interaction	2.34***	3.17***	2.76	2.48***
	(0.91, 3.77)	(1.03, 5.31)	(-1.29, 6.81)	(1.34, 3.62)
Constant	47.19***	47.72***	52.63***	47.91***
	(44.85, 49.54)	(44.10, 51.34)	(46.11, 59.15)	(46.00, 49.82)
Observations	1,081	462	309	1,852
Log Likelihood	-4,466.12	-1,856.76	-1,320.73	-7,657.56

Note. The first three columns show consistency in parameter estimates across three separate cohorts of intellectually precocious youth; the last column gives the best parameter estimates for the model fit with all cohorts combined. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

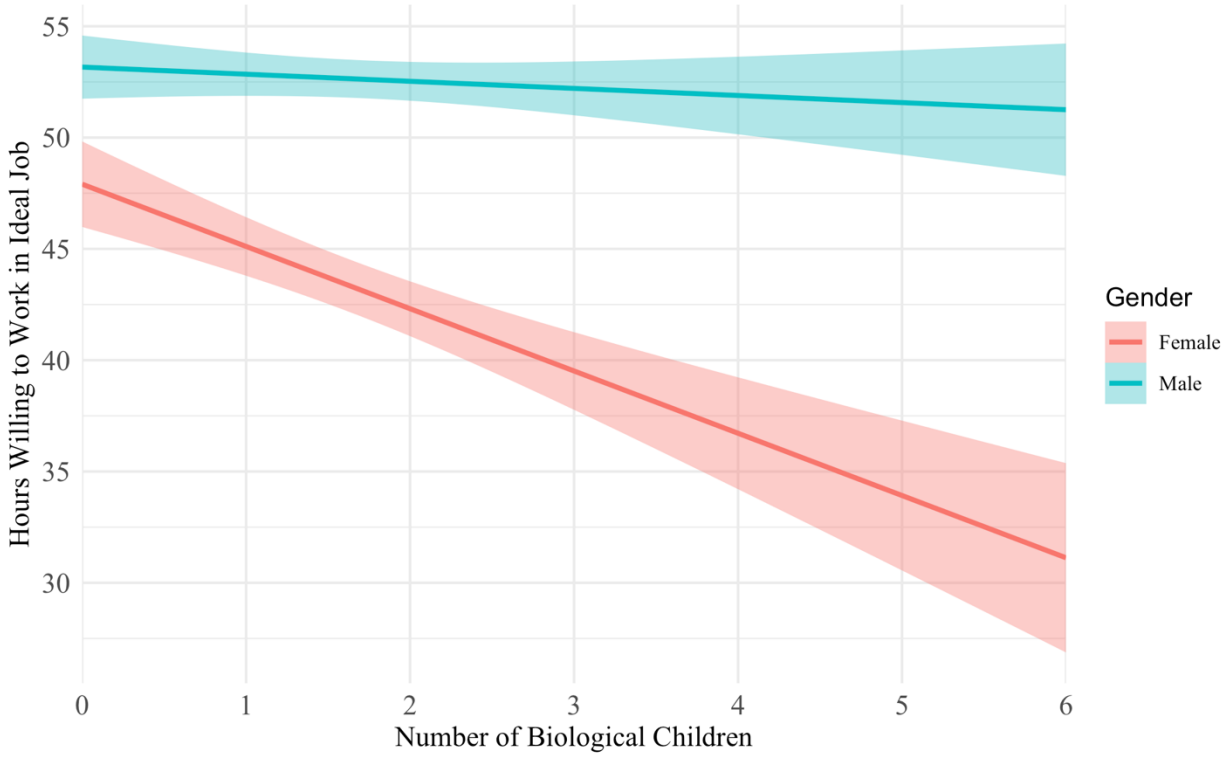

Figure 9. Study 1: Hours Willing to Work as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Subjective Well-Being

Parallel analysis (Horn, 1965) was conducted to derive an ideal linear combination of five indicators of psychological well-being. Figure 10 shows the results using Revelle's (2021) psych package. Only the first eigenvalue of the correlation matrix of the psychological well-being measures was above what would be expected by chance. Thus, principal component analysis was conducted and the first component was retained (see Table 7 for a summary of the principal component structure). Results mirrored those of Bernstein and colleagues (2021), which used a similar sample and analytic procedure (but see also Lucas and colleagues, 1996).

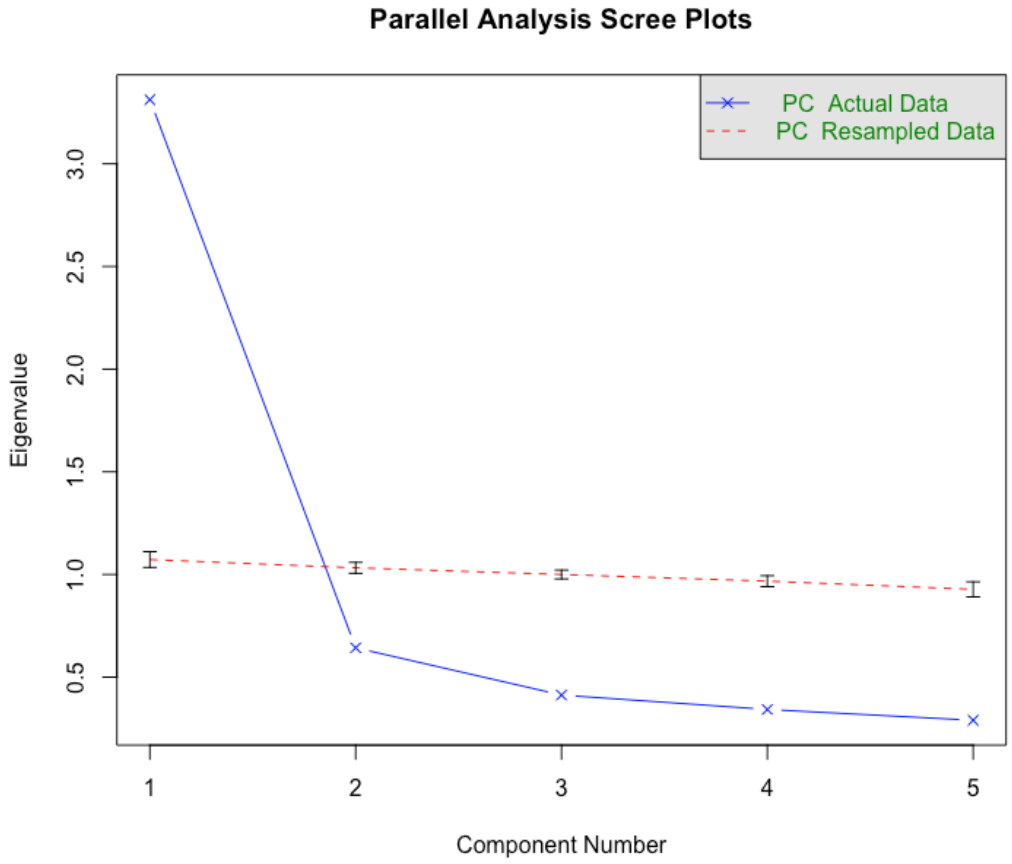

Figure 10. Scree Plot and Parallel Analysis for Talent Search Participants.
Performed on the correlation matrix of the indicators of psychological well-being for Study 1, only the first eigenvalue is above what would be expected by chance.

Table 7: Principal Component Analysis Results for Talent Search Participants

	Component Loading	Communality	Uniqueness
Core Self-Evaluations	0.85	0.73	0.27
Positive Affect	0.81	0.66	0.34
Life Satisfaction	0.83	0.68	0.32
Negative Affect (Reversed)	0.74	0.55	0.45
Psychological Flourishing	0.83	0.69	0.31

Note. All indicators had relatively strong and uniform loadings on the first principal component, suggesting relatively equal weighting in its computation.

Table 8 summarizes the regression results of the psychological well-being principal component as a function of number of number of children by gender; Figure 11 shows the relations graphically. Increases in psychological well-being as a function of number of children was observed for women (simple slope = .139, $t(1644) = 4.381$, $p < .001$). Moreover, the interaction term was virtually zero *(*$b = 0.007$, $t(1644) = 0.181$, $p = .856$), resulting in commensurate increases in life satisfaction as a function of number of children for men (simple slope = .147, $t(1644) = 6.196$, $p < .001$).

Table 8: Modeling Psychological Well-Being for Talent Search Participants

	Dependent variable:			
	Psychological Well-Being			
	Cohort 1	Cohort 2	Cohort 3	Talent Search
Number of Biological Children	0.11***	0.22***	0.13	0.14***
	(0.05, 0.18)	(0.07, 0.37)	(-0.10, 0.36)	(0.08, 0.20)
Gender	-0.02	-0.05	-0.08	-0.07
	(-0.21, 0.17)	(-0.39, 0.30)	(-0.57, 0.41)	(-0.23, 0.09)
Interaction	0.03	-0.04	-0.04	0.01
	(-0.06, 0.12)	(-0.22, 0.13)	(-0.30, 0.23)	(-0.07, 0.08)
Constant	-0.12	-0.29*	-0.33	-0.18***
	(-0.27, 0.02)	(-0.58, 0.01)	(-0.75, 0.08)	(-0.31, -0.05)
Observations	944	420	284	1,648
Log Likelihood	-1,244.39	-599.43	-429.84	-2,295.64

Note. The first three columns show consistency in parameter estimates across three separate cohorts of intellectually precocious youth; the last column gives the best parameter estimates for the model fit with all cohorts combined. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

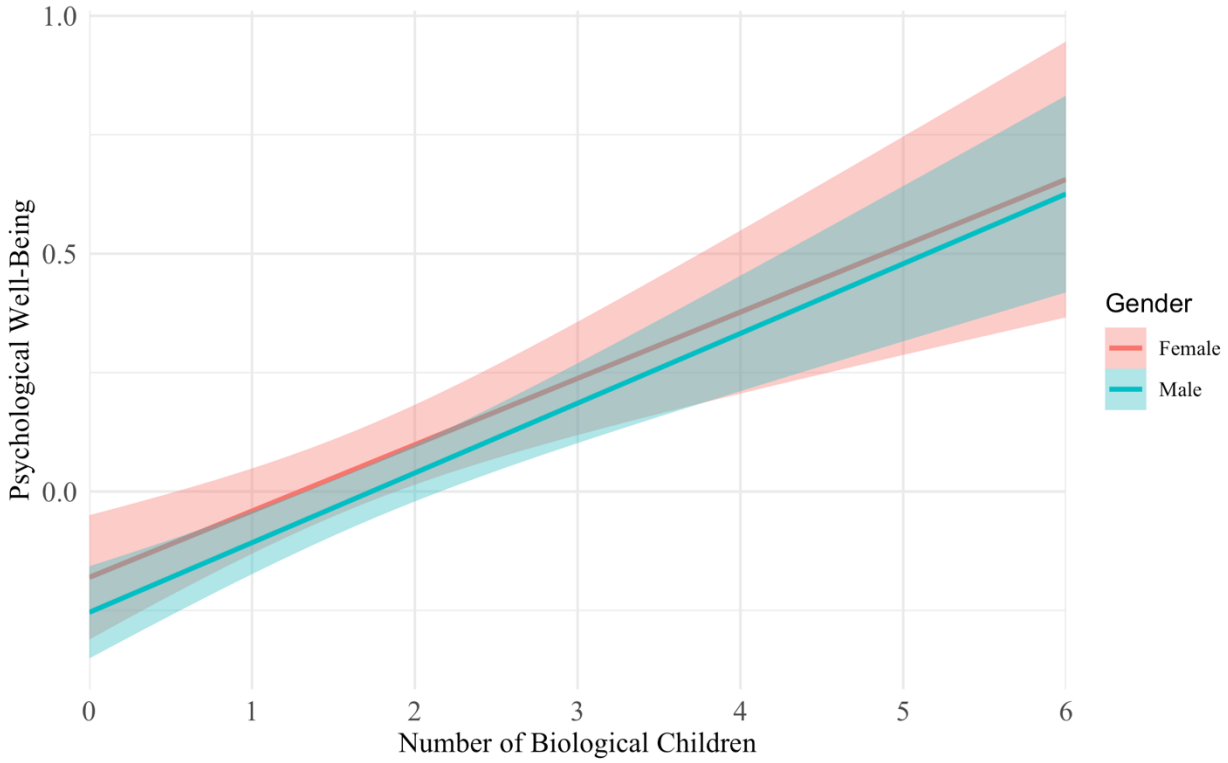

Figure 11. Study 1: Psychological Well-Being as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Figure 12 summarizes mean levels of relationship satisfaction faceted by number of children and gender. There was very little variation in this one-item measure; most participants who were in relationships described themselves as either satisfied or very satisfied. As a result, it was difficult to test whether men and women with more children were more satisfied in their relationships than those with fewer or no children. Nevertheless, it is clear (and psychologically important) that participants in relationships were generally satisfied, suggesting again many different ways in which to live a fulfilling and satisfying life for high-potential individuals.

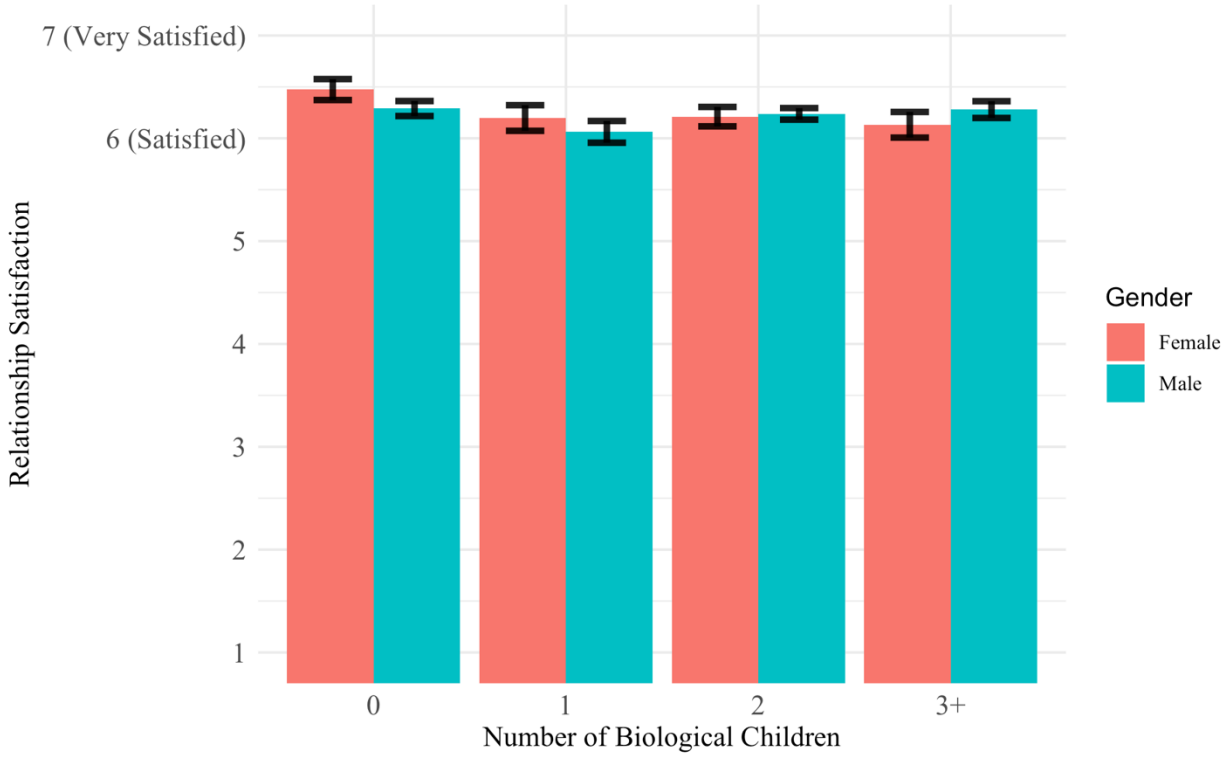

Figure 12. Study 1: Relationship Satisfaction as a Function of Number of Children. Error bars represent ± 1 standard error of the mean.

Table 9 summarizes the regression results of life satisfaction as a function of number of children by gender; Figure 13 shows the relations graphically. Women reported increases in life satisfaction as a function of having more children (simple slope = 1.026, $t(1757) = 5.177$, $p < .001$). Moreover, the interaction term was virtually zero ($b = 0.040$, $t(1757) = 0.161$, $p = .872$), resulting in commensurate increases in life satisfaction as a function of number of children for men (simple slope = 1.066, $t(1757) = 7.223$, $p < .001$).

Table 9: Modeling Life Satisfaction for Talent Search Participants

	Dependent variable:			
	Life Satisfaction			
	Cohort 1	Cohort 2	Cohort 3	Talent Search
Number of Biological Children	1.00*** (0.56, 1.45)	1.32*** (0.40, 2.24)	0.89 (-0.49, 2.28)	1.03*** (0.64, 1.41)
Gender	-0.43 (-1.69, 0.82)	-0.66 (-2.76, 1.43)	-2.32 (-5.26, 0.61)	-0.81 (-1.81, 0.19)
Interaction	0.12 (-0.47, 0.71)	-0.28 (-1.35, 0.78)	0.05 (-1.51, 1.61)	0.04 (-0.44, 0.52)
Constant	24.25*** (23.30, 25.20)	24.38*** (22.62, 26.15)	25.59*** (23.06, 28.11)	24.48*** (23.68, 25.28)
Observations	1,018	441	302	1,761
Log Likelihood	-3,288.96	-1,433.44	-1,002.90	-5,728.37

Note. The first three columns show consistency in parameter estimates across three separate cohorts of intellectually precocious youth; the last column gives the best parameter estimates for the model fit with all cohorts combined. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

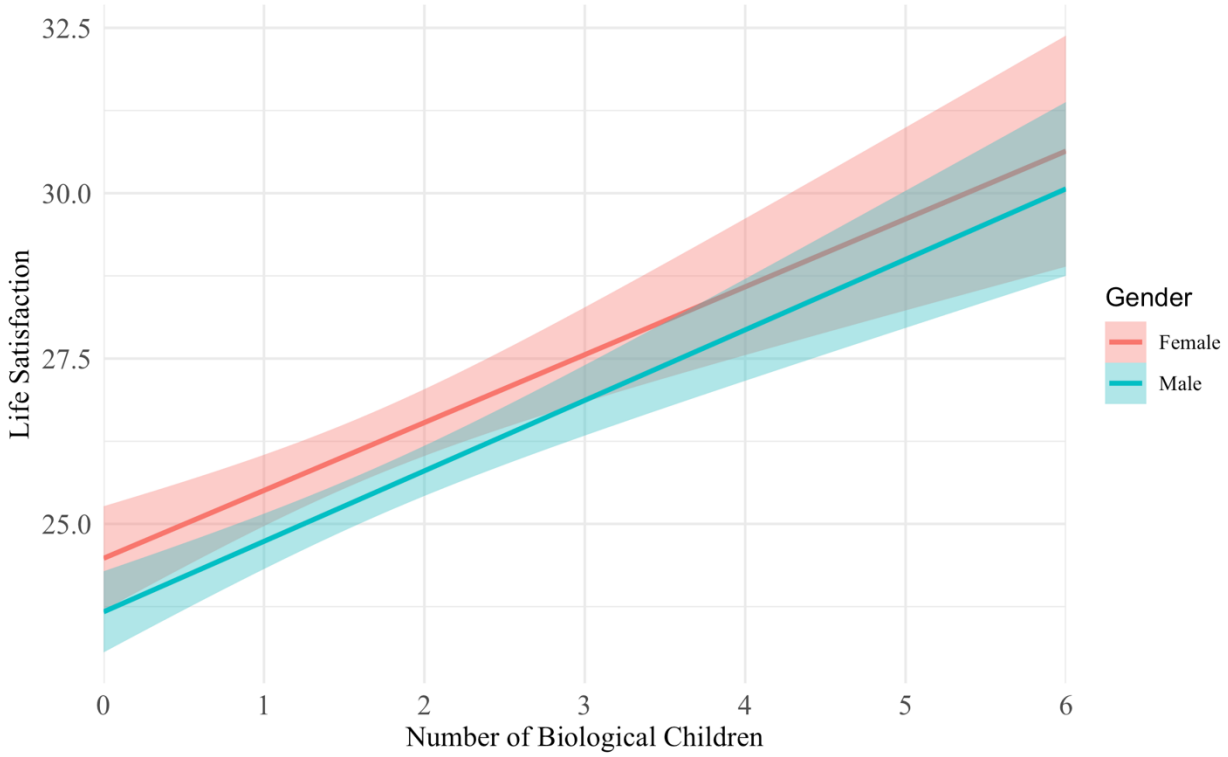

Figure 13. Study 1: Life Satisfaction as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Ancillary Analyses

After examining the results in Study 1, two additional analyses seemed to have potential to inform the set of hypotheses used to frame this study. That high-potential men and women's occupational outcomes and work preferences diverge so substantially with parenthood and increasing children—and that their subjective well-being both improve with more children—necessitates asking whether they are "simply" finding alternative ways in which to find meaning and satisfaction in their lives. Successful work teams divide the labor of the tasks that they need to accomplish, and partners in a relationship are no different. Given this, I decided to analyze data involving: 1. Partner's full-time employment and 2. Total family income.

Figure 14A reveals that as the number of children for men increases, the probability of having a significant other working full time decreases; conversely, as female participants have more children, their partners are more likely to work full time. This pattern suggests a division of labor that partially reconciles the previous findings of divergence in income for mothers and fathers as a function of children coinciding with increases in subjective well-being for both: mothers and fathers simply adapt to changes in demands at home in different ways. How does this relate to total family income?

Figure 14B shows data for total family income (e.g., participant and spouse income and investment income) as a function of number of children. This finding suggests that a division of labor between partners does not imply decreases in economic returns. These ancillary findings can help to explain the finding that, although male and female participants' individual incomes were differentially affected by parenthood and the number of children they have, their total family income remained relatively high while their reports of subjective well-being increased with more children. They do not appear to be economically harmed and most, presumably, wanted more children and were happy they had them.

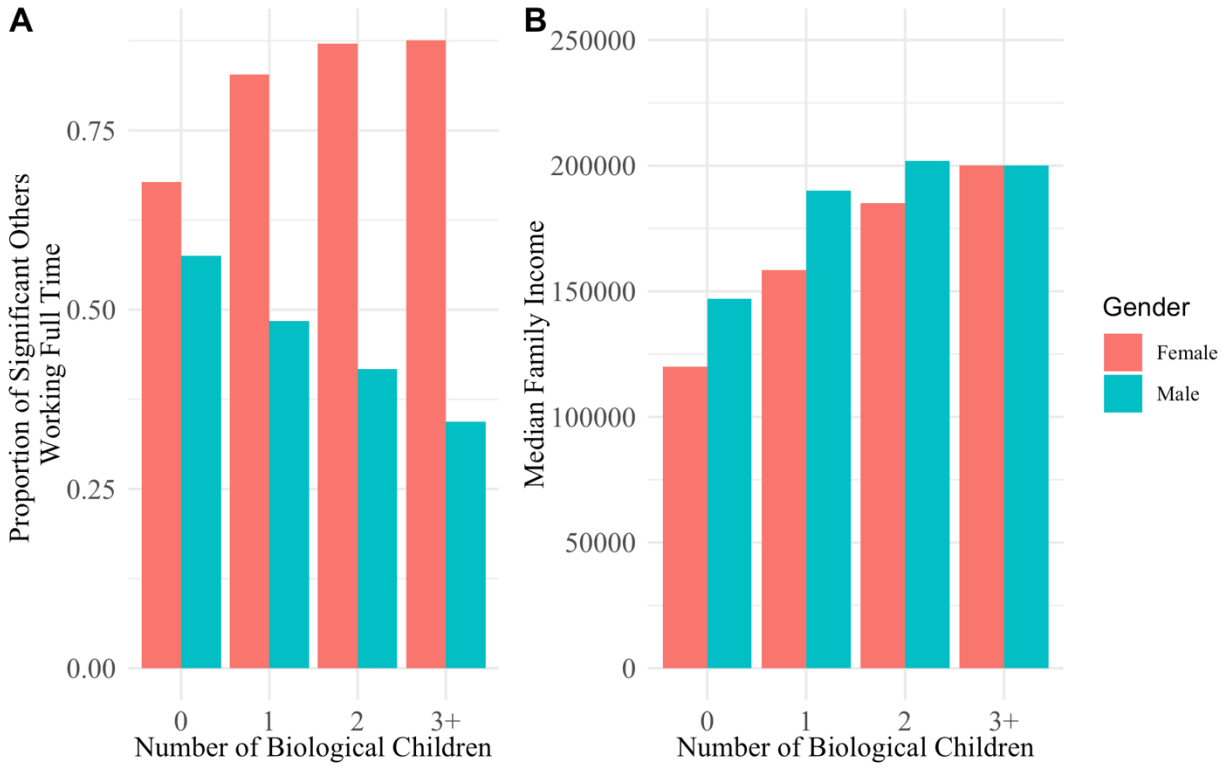

Figure 14. Division of Labor and Family Income amongst Talent Search Participants. Part A shows diverging trends in the proportion of male and female participants' significant others working full time as a function of number of children. Part B shows an increase in the median family income as a function of number of children.

Discussion

Across three cohorts of intellectually prodigious participants, the results from Study 1 replicate robust normative findings on the income divergence between women and men as a function of the number of children they have. Consistent with the literature on the Motherhood Penalty and Fatherhood Bonus (Cukrowska-Torzewska & Matysiak, 2020), intellectually talented females had substantially lower incomes as a function of the number of children that they had (i.e., Motherhood Penalty); on the other hand, intellectually talented male participants experience a complete reversal of this relationship and earned substantially larger incomes with more children (i.e., Fatherhood Bonus). These findings support hypothesis H_{A1}, that the incomes

of men and women would be differentially affected by the presence of children. The results also are consistent with what is predicted by the extensive literature on the Motherhood Penalty (Budig & England, 2001).

However, other outcomes examined in the context of these dynamic income changes are psychologically informative. They highlight why it is important to take a broader lifespan developmental perspective when modeling important life outcomes in general and income disparities between the genders in particular.

The decision to study income changes as a function of parenthood in the context of concomitant outcomes such as participants' relationship satisfaction, psychological well-being, total family income, and overall life satisfaction was illuminating. While the Fatherhood Bonus and Motherhood Penalty appear to be genuine and robust psychological phenomena (even at the outer edge of the envelope of extraordinary human capital), other valued aspects of life that co-occur with these changes either remain uniformly high or increase. As income changes diverged between the genders as a function of parenthood and number of children, both genders experienced overall increases in their psychological well-being and life satisfaction. In addition, their relationship satisfaction remained the same (uniformly high) and their total median family income for participants with children was above $150,000 across one to three or more children. Indeed, as the partners of female participants worked more, and the partners of male participants worked less, there was a trend for participants with children to have significantly *more* total family income relative to households without children, $r(1559) = .15$, $p < .001$.

The results from Study 1 paint a fuller and somewhat more positive picture of the ways in which intellectually prodigious participants structure their lives. Importantly, these are highly able participants with the personal resources and opportunities to do many things. They are likely

highly sought after in the workplace. They likely experience many opportunities to implement and develop their talents in ways that are highly compensated. It is fair to say that they are likely to have more possibilities and opportunities than the average person.

The prediction of log-odds of leadership and probability of leadership also both suggested an interactive effect by which men with more children experienced higher likelihoods of occupational leadership. The results from this analysis do not technically support H_{A2} (which stated that mothers also would see a decrement to likelihood of leadership), although they are consistent with the hypothesized theme (i.e., interactive effects of children on occupational outcomes for mothers and fathers). Several explanations for this inconsistency seem plausible, including the fact that the base rate of occupational leadership among women and men in the sample was, by definition, low. It is possible that a true decrease in probability of occupational leadership for women with more children in this sample was statistically unidentifiable, given the low overall base rates of occupational leadership. Intense floor effects on this metric may have compromised this analysis.

The work-life structure results also confirm H_{B1}, H_{B2}, and H_{B3}. That is, high-potential female participants would elect to spend more time with family and the home (as compared to at work) as more children come along; and high-potential male participants would elect to spend more time at work (as compared to with family and the home) as more children come along. Female participants with more children reported spending substantially fewer hours at work and substantially more hours with their families and homes than did women with few or no children; they also reported a desire to work fewer hours, even if given their ideal job. Fathers, on the other hand, reported no changes in the number of hours worked, slight increases in the number of

hours spent with the family, and no changes in the number of hours willing to work if given their ideal job.

Finally, to specifically address other hypothesized relationships statistically, both men and women who chose to have children derived significantly more psychological well-being and life satisfaction from doing so, providing positive evidence for H_{C1} and H_{C3} (that men and women who have children do seem to derive greater psychological well-being and life satisfaction from doing so). H_{C2} (that both men and women would derive greater relationship satisfaction from more children) could not be meaningfully tested directly due to lack of variability to be predicted in relationship satisfaction (it was uniformly high, regardless of how many children they had).

The ancillary results help tie together the seemingly incompatible findings that both high-potential men and women with more children derived greater subjective well-being while their work-life structures changed in different ways with women experiencing decreases in their occupational achievement. The significant others of the intellectually prodigious male participants tended to work less as more children came along, and the reverse was true for the significant others of the intellectually prodigious females. Nevertheless, participants and their significant others tended to earn more total family income as more children came along.

These findings suggest a gender-differentiating division of labor tendency between participants and their significant others in which one partner takes on more responsibility in earning income (tangible support for the family) while the other partner takes on more responsibility with hands-on family support. Importantly, the net result of this division of labor is *not* a decrease to the economic resources for the family unit. Rather, it seems that the partnership generally works well across multiple outcomes and leads to increases in overall subjective well-being and life satisfaction while romantic relationship satisfaction remains uniformly high. But

will these findings hold up in other "types" of high-potential participants? For example, will these findings be found among world-class STEM doctoral students with gender-commensurate psychological profiles across specific cognitive abilities, educational-occupational interests, and values orientation, plus comparably rich and STEM-concentrated educational histories?

Study 2

Study 1 was designed as a series of operational replications across three cohorts of successively more able intellectually prodigious participants. Study 2 was designed as a constructive replication to examine the extent to which findings observed in Study 1 would mirror those obtained among world-class STEM doctoral students trained in universities ranked among the best in the world. Given the contemporary emphasis of developing STEM leaders (Ceci et al., 2021; National Science Foundation, 2021; Stewart-Williams & Halsey, 2021), this generalization probe seems especially timely.

Constructive replications are built around the idea of varying as many irrelevant design features of an initial study as possible while maintaining focus on the operative constructs of interest. Top STEM doctoral students, like the participants in Study 1, are clearly high potential; however, they differ psychologically from the participants in Study 1 in many important respects. For example, characteristic gender differences found in intellectually prodigious populations are not observed. That is, among top STEM doctoral students, the genders display highly uniform ability, interests, and values profiles; furthermore, their educational histories are STEM-concentrated well before high school graduation and in college (Lubinski et al., 2001).

Specifically, among elite STEM doctorial students, men and women displayed high levels of mathematical ability (relative to verbal ability), dominant scientific interests, regnant theoretical values (relatively lower religious values), a preference for math and science courses in high school, and a propensity to work long hours. When compared to mathematically prodigious individuals more generally, top STEM doctoral students are psychologically much more similar to mathematically precocious *males* rather than mathematically precocious *females*. To be sure, mathematically precocious females are just as impressive, but their mathematical/verbal ability pattern is much more balanced and their interests and values much more eclectic than either their male counterparts or top STEM doctoral students (Lubinski et al., 2001; K. O. McCabe et al., 2020). Mathematically prodigious populations also display characteristic gender differences in the people-versus-things (or inorganic versus organic) occupations that they ultimately secure (Su et al., 2009).

Given the psychological similarities across both genders among these elite doctoral students, they are an intriguing sample to address Canizares's (Mervis, 2009) earlier noted suggestion for studying STEM doctoral students for understanding differential developmental trajectories across the STEM *occupational pipeline*. Several of the gender differentiating ability and interest patterns that lead to contrasting career paths (and even their educational histories) are controlled and, additionally, their levels of talent and STEM graduate training are both world class. Motivating this study further is Cech and Blair-Loy's (2019) observation that new mothers were far less likely than similarly qualified new fathers to earn tenure-track professorships and, in addition to be promoted once in those jobs (Mason et al., 2013). Yet, we do not have a good sense of the overall longitudinal gender-related impact of parenthood on top STEM professionals. Therefore, a constructive replication of Study 1 involving elite STEM doctoral

seems especially well situated to be particularly informative. A sample of these participants will be studied here to test all of the hypotheses examined in Study 1.

Methods

Participants

The cohort of top STEM doctoral students was comprised of individuals who were identified in 1992 as first- or second-year doctoral students pursuing training at one of the top 15 STEM doctoral programs in the United States (Gourman, 1989; Lubinski et al., 2001; National Research Council, 1987). Their average age at the time of identification was 25; women were oversampled to obtain comparable numbers of men and women for analytic purposes. This sample consisted of 255 females and 267 males (n = 522); the sample was 84.3% white or Caucasian, 1.2% black or African American, 1.9% Hispanic, 8.4% Asian, Asian-American, or Pacific Islander, and 4.2% Other. This cohort's response rate to the mid-career survey was 77.1%.

Procedure, Instrumentation, and Analytical Approach

The top STEM doctoral students completed the mid-career survey between 2017 and 2018. They were assessed simultaneously with Study 1's Cohort 3 participants and under the same procedures. The same information was collected on these participants: occupational status, gender identification, number of children, preferences for work structure, relationship satisfaction, and five scales of psychological well-being. As in Study 1, due to sparsity beyond six biological children (see Figure 15), all categories greater than or equal to six biological children were collapsed into six children. The alpha reliabilities of Core Self-Evaluations, Positive Affect, Life Satisfaction, Negative Affect (reversed), and Psychological Flourishing

were .86, .90, .88, .88, and .83, respectively. Additionally, occupational leadership was coded during criterion searches conducted in 2016 and 2017. The same statistical framework and computing packages were used in Study 2.

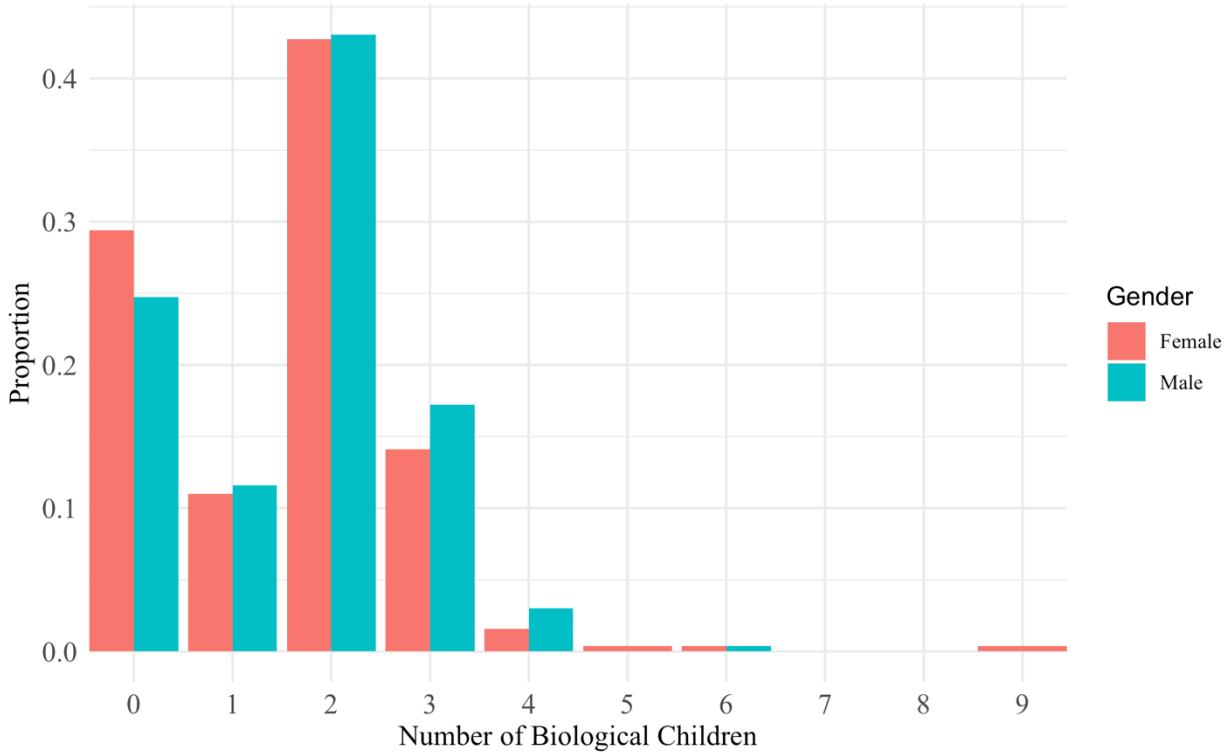

Figure 15. Distribution of Biological Children for Top STEM Doctoral Students.

Results

Occupational Outcomes

Table 10 summarizes the regression results of log wages as a function of number of children by gender; results are juxtaposed with the parameter estimates from the talent search participants in Study 1. Figure 16 shows the relationship graphically. Parameter estimates were consistent with Study 1: greater numbers of children were associated with lower log wages for women (simple slope = -0.201, $t(493) = -1.638$, $p = .102$), yet a positive and strong interaction

reversed the trend for men (simple slope = 0.224, $t(493) = 1.801$, $p = .072$). The interaction term was significant and in the same direction as the parameter estimate in Study 1 ($b = 0.425$, $t(493) = 2.432$, $p = .015$).

Table 10: Modeling Log Participant Income for Top STEM Doctoral Students

	Dependent variable:				
	Log Participant Income				
	Cohort 1	Cohort 2	Cohort 3	Talent Search	Top STEM Graduate Students
Number of Biological Children	-0.41*** (-0.62, -0.21)	-0.86*** (-1.24, -0.48)	-0.25 (-0.86, 0.36)	-0.48*** (-0.66, -0.31)	-0.20 (-0.44, 0.04)
Gender	0.03 (-0.54, 0.60)	-0.24 (-1.14, 0.66)	0.34 (-0.93, 1.61)	0.03 (-0.42, 0.47)	0.25 (-0.43, 0.93)
Interaction	0.65*** (0.38, 0.92)	1.19*** (0.74, 1.64)	0.55 (-0.13, 1.24)	0.75*** (0.54, 0.97)	0.43** (0.08, 0.77)
Constant	10.92*** (10.48, 11.36)	11.12*** (10.37, 11.87)	10.60*** (9.50, 11.70)	10.91*** (10.55, 11.26)	11.14*** (10.68, 11.61)
Observations	1,071	461	316	1,848	497
Log Likelihood	-2,640.04	-1,130.18	-797.02	-4,570.66	-1,114.10

Note. Consistency is shown between Study 1 and Study 2. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

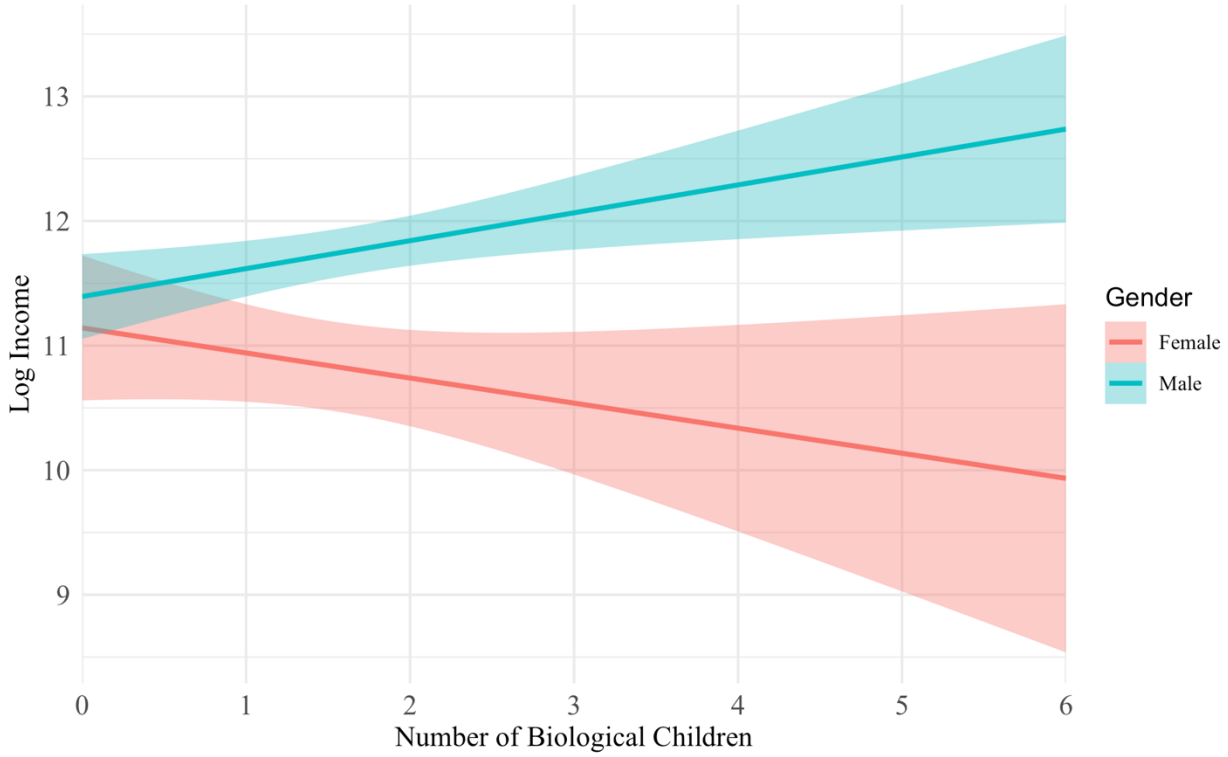

Figure 16. Study 2: Log Income as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Figure 17 shows a graphical representation of the results of the same model on the natural scale of interpretation: absolute income. On this scale, a one-unit increase in the number of children for women is associated with an 18% decrease in wages, whereas the same change is increased with a 26% increase in wages for men. The result is a distinct separation in income between men and women as a function of the number of children they have.

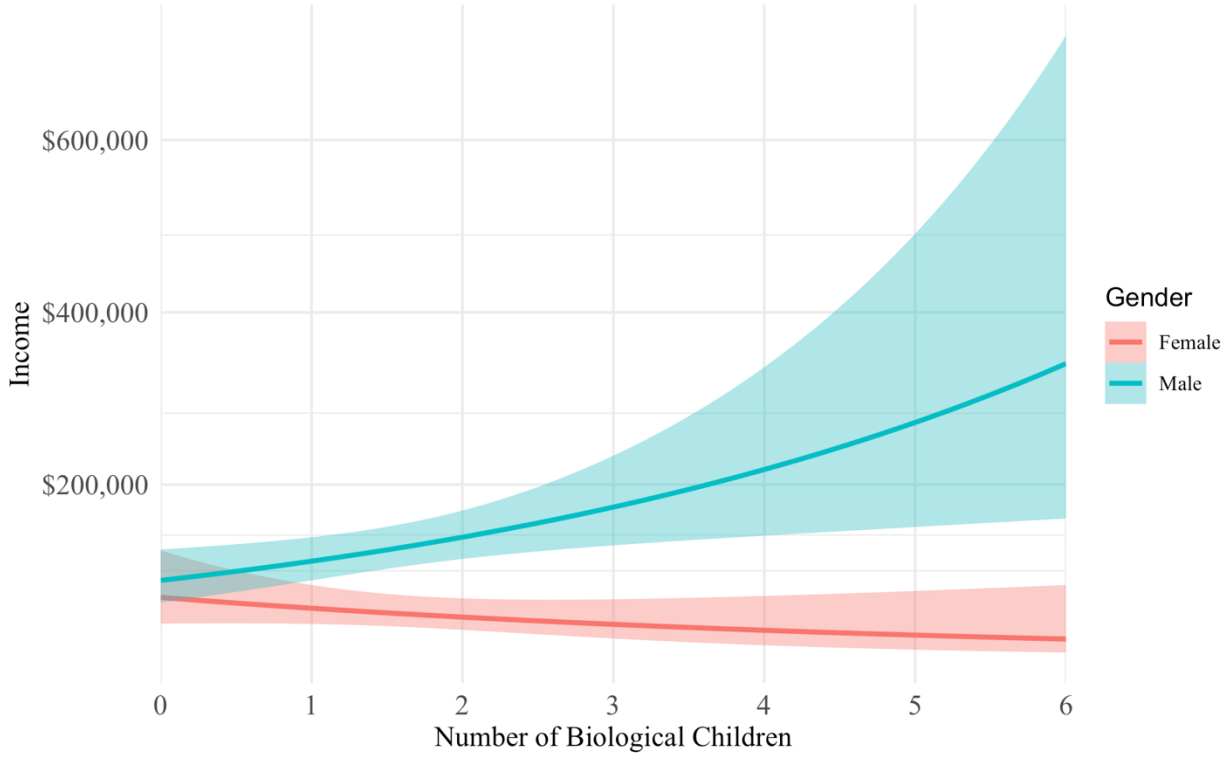

Figure 17. Study 2: Income as a Function of Number of Children.
Shaded bands give 95% confidence intervals about the exponentiated regression line.

Table 11 summarizes the regression results for the log odds of leadership as a function of number of children by gender. Again, results are juxtaposed with the parameter estimates from the talent search participants in Study 1. Figure 18 shows the relations graphically. Parameter estimates were consistent with Study 1: number of children is related to lower log-odds of leadership for women (simple slope = -0.024, $t(518) = -0.190$, $p = .849$); yet, a positive and strong interaction ($b = 0.208$, $t(518) = 1.219$, $p = .223$) reversed the trend for men (simple slope = 0.184, $t(518) = 1.608$, $p = .107$). Reflecting the smaller sample size of Study 2, the interaction term was not significant, but it was consistent with (and nearly identical to) the parameter estimates in Study 1.

Table 11: Modeling Log Odds of Leadership for Top STEM Doctoral Students

	Dependent variable:				
	Leadership				
	Cohort 1	Cohort 2	Cohort 3	Talent Search	Top STEM Graduate Students
Number of Biological Children	-0.08 (-0.34, 0.18)	0.14 (-0.25, 0.53)	-0.07 (-0.55, 0.42)	-0.05 (-0.24, 0.14)	-0.02 (-0.27, 0.22)
Gender	0.51 (-0.11, 1.14)	0.60 (-0.32, 1.52)	-0.19 (-1.17, 0.79)	0.44* (-0.01, 0.88)	0.21 (-0.45, 0.87)
Interaction	0.27* (-0.03, 0.56)	-0.06 (-0.50, 0.38)	0.31 (-0.23, 0.84)	0.21* (-0.01, 0.43)	0.21 (-0.13, 0.54)
Constant	-2.27*** (-2.80, -1.75)	-2.12*** (-2.93, -1.32)	-1.15*** (-2.00, -0.29)	-2.02*** (-2.40, -1.64)	-1.21*** (-1.68, -0.73)
Observations	1,134	487	331	1,952	522
Log Likelihood	-462.01	-226.05	-190.90	-890.86	-304.11

Note. Consistency is shown between Study 1 and Study 2. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

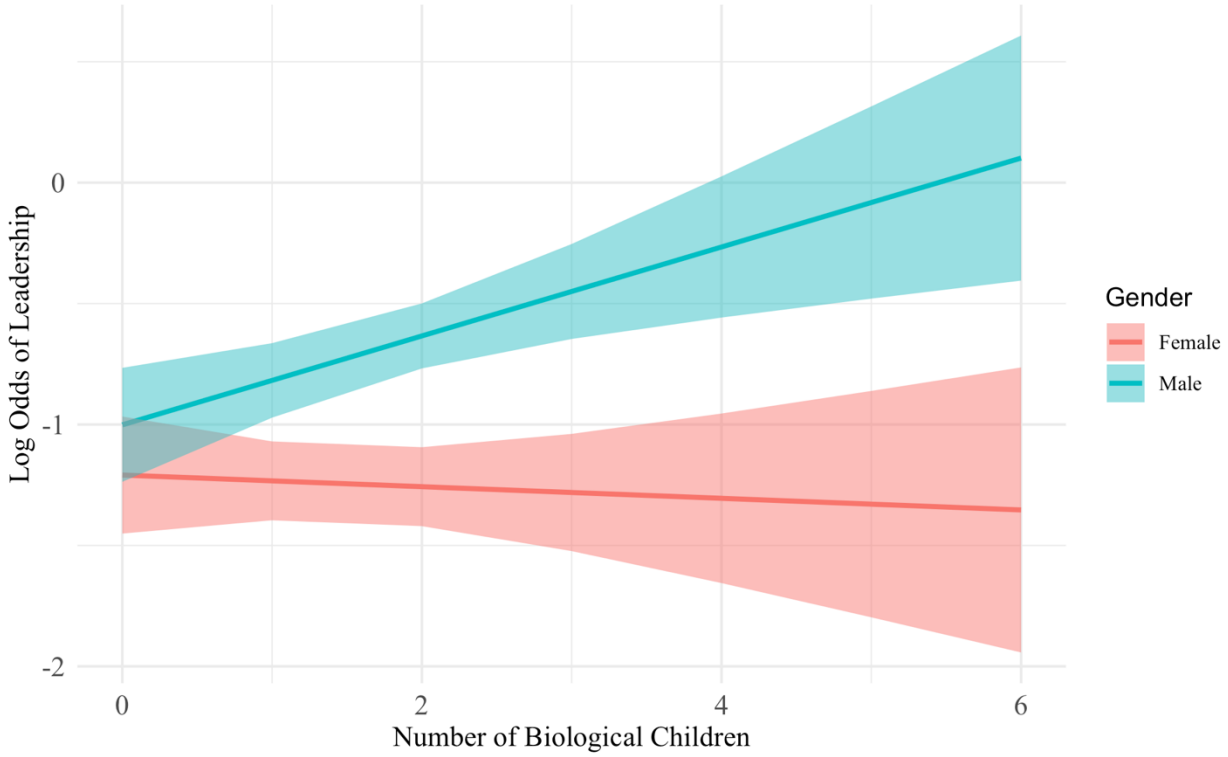

Figure 18. Study 2: Log Odds of Leadership as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Table 12 summarizes the interaction of gender and the number of children on the probability of leadership; Figure 19 presents the model graphically. The interaction is not statistically significant at any value of number of children, although the interaction becomes larger as the number of children increases. Nevertheless, the magnitude of the interaction is consistent with the results in Study 1 as well as with the results based on the log odds scale: more children were associated with larger increases in probability of leadership for men compared to women.

Table 12: Interaction Estimates in Probability of Leadership for Top STEM Doctoral Students

# Biological Children	Interaction Estimate	Standard Error of Interaction Estimate	Statistically Significant?
0	.0404	.0299	No
1	.0433	.0321	No
2	.0458	.0341	No
3	.0478	.0356	No
4	.0492	.0361	No
5	.0499	.0354	No
6	.0498	.0335	No

Note. The interaction can be interpreted as the increased marginal effect of the number of biological children on probability of leadership for a change in gender from female to male.

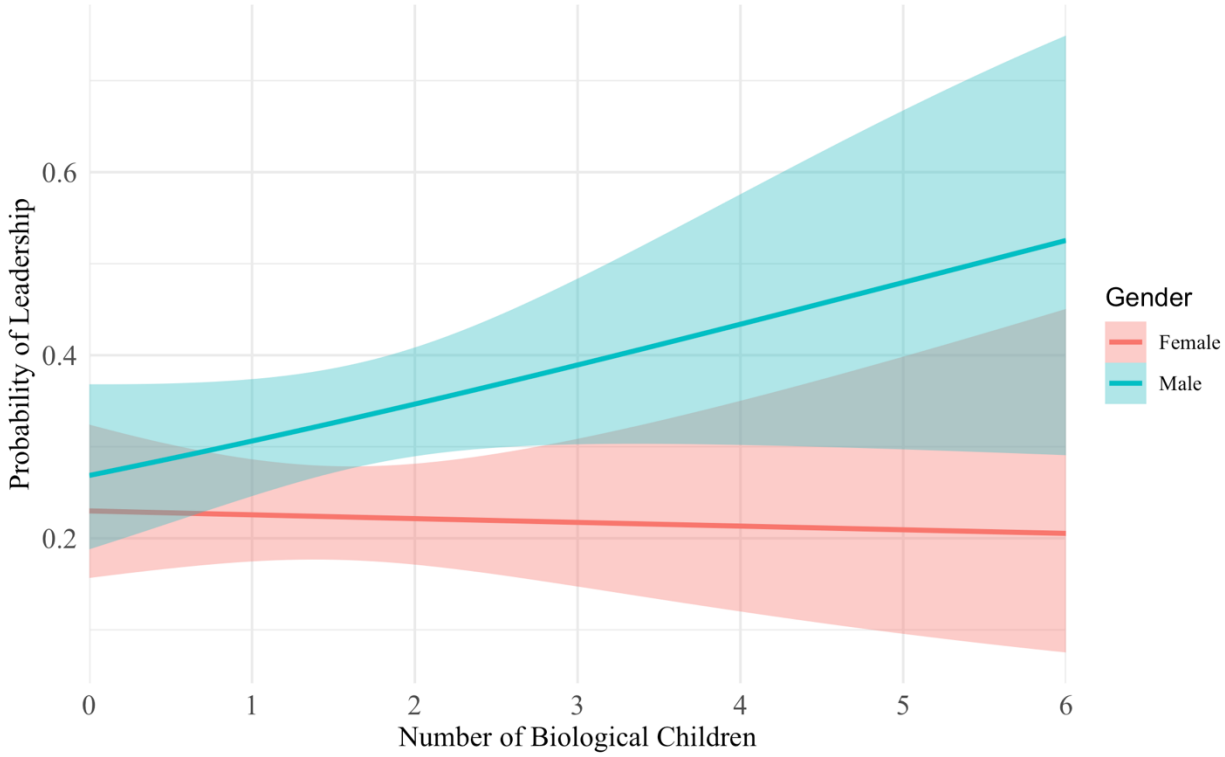

Figure 19. Study 2: Probability of Leadership as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Work-Life Structure

Table 13 summarizes the regression results of hours worked as a function of number of children by gender; Figure 20 shows the relations graphically. Parameter estimates were consistent with Study 1: more children were associated with lower hours worked for women (simple slope = -2.040, $t(447) = -2.331$, $p = .020$), and a positive and strong interaction made null the trend for men (simple slope = 0.811, $t(447) = 0.884$, $p = .377$). The interaction term was significant and in the same direction as the parameter estimate in Study 1 ($b = 2.851$, $t(447) = 2.249$, $p = .025$).

Table 13: Modeling Hours Worked for Top STEM Doctoral Students

	Dependent variable:				
	Hours Working				
	Cohort 1	Cohort 2	Cohort 3	Talent Search	Top STEM Graduate Students
Number of Biological Children	-3.08*** (-4.31, -1.85)	-5.42*** (-7.39, -3.46)	-2.94 (-6.58, 0.71)	-3.52*** (-4.53, -2.52)	-2.04** (-3.76, -0.32)
Gender	2.29 (-1.08, 5.66)	-0.42 (-4.99, 4.15)	-3.63 (-11.24, 3.98)	0.50 (-2.06, 3.06)	-0.79 (-5.72, 4.13)
Interaction	2.88*** (1.29, 4.47)	6.51*** (4.24, 8.79)	4.42** (0.30, 8.54)	4.00*** (2.76, 5.24)	2.85** (0.37, 5.34)
Constant	46.72*** (44.15, 49.30)	46.60*** (42.77, 50.43)	46.73*** (40.20, 53.26)	46.56*** (44.52, 48.61)	47.19*** (43.82, 50.55)
Observations	958	425	267	1,650	451
Log Likelihood	-4,002.15	-1,707.80	-1,128.81	-6,855.97	-1,878.43

Note. Consistency is shown between Study 1 and Study 2. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

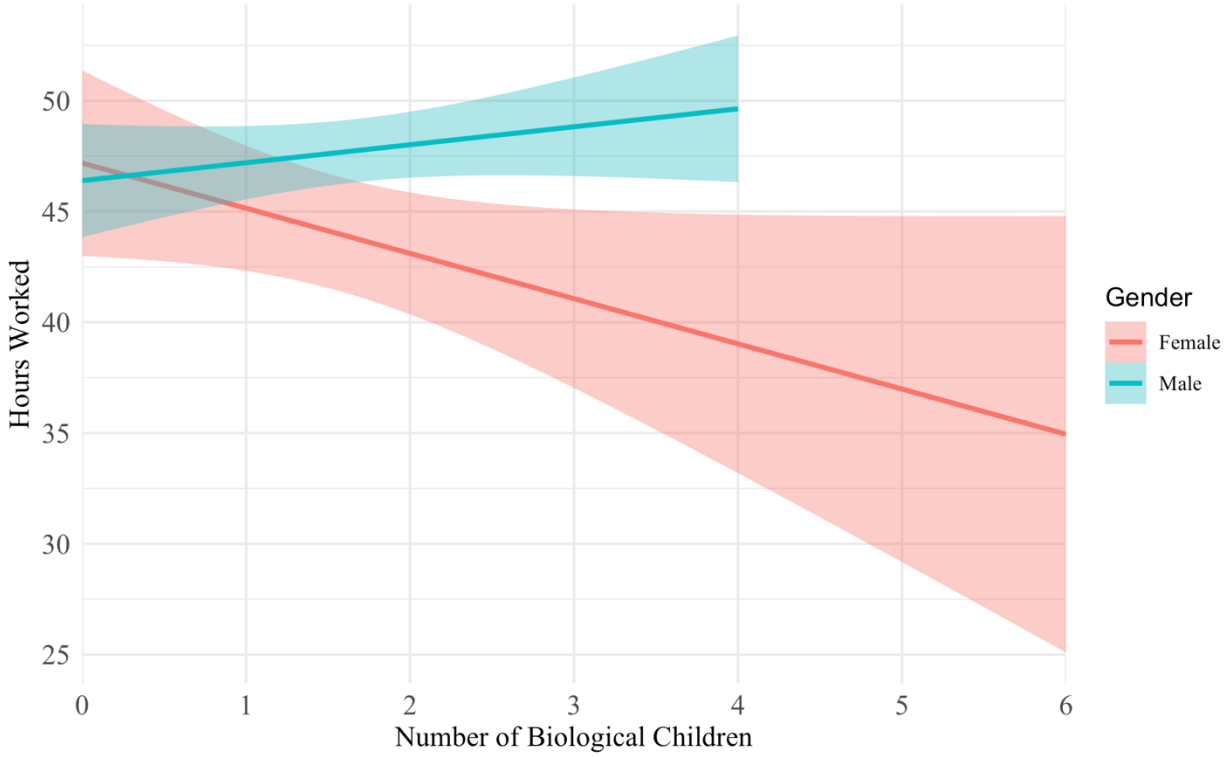

Figure 20. Study 2: Hours Worked as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Table 14 summarizes the regression results of hours spent with family and home maintenance as a function of number of children by gender; Figure 21 shows the relations graphically. Parameter estimates were consistent with Study 1: more children were associated with more hours spent with the family and home for women (simple slope = 6.536, $t(436) = 4.260$, $p < .001$), and a significant negative interaction ($b = -5.179$, $t(436) = -2.344$, $p = .020$) attenuated the trend for men (simple slope = 1.356, $t(436) = 0.853$, $p = .394$). The interaction term was in the same direction and of similar magnitude as the parameter estimate in Study 1.

Table 14. Modeling Hours with Family and Home for Top STEM Doctoral Students

	Dependent variable:				
	Hours with Family				
	Cohort 1	Cohort 2	Cohort 3	Talent Search	Top STEM Graduate Students
Number of Biological Children	5.12*** (2.95, 7.30)	17.54*** (13.72, 21.36)	7.11*** (2.26, 11.95)	7.99*** (6.22, 9.75)	6.54*** (3.53, 9.54)
Male	-1.73 (-7.71, 4.24)	7.38 (-1.55, 16.31)	2.54 (-7.33, 12.40)	-0.30 (-4.80, 4.19)	-1.06 (-9.69, 7.57)
Interaction	-3.29** (-6.10, -0.47)	-15.33*** (-19.77, -10.89)	-4.65* (-10.11, 0.80)	-5.88*** (-8.05, -3.71)	-5.18** (-9.51, -0.85)
Constant	41.40*** (36.81, 45.98)	27.74*** (20.28, 35.19)	30.68*** (22.16, 39.19)	37.42*** (33.82, 41.02)	41.45*** (35.52, 47.39)
Observations	929	417	252	1,598	440
Log Likelihood	-4,391.43	-1,953.25	-1,120.52	-7,502.78	-2,070.66

Note. Consistency is shown between Study 1 and Study 2. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

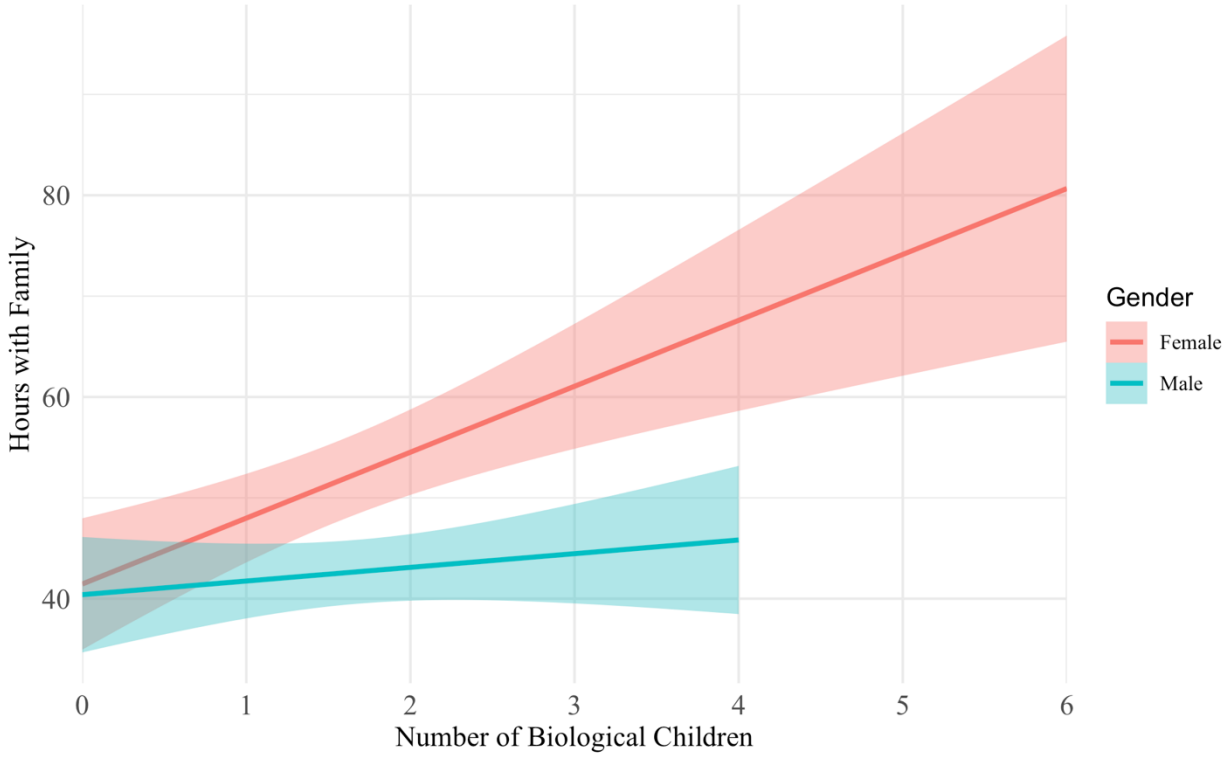

Figure 21. Study 2: Hours with Family as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Table 15 summarizes the regression results of hours willing to work as a function of number of children by gender; Figure 22 shows the relations graphically. Parameter estimates were consistent with Study 1: more children were associated with fewer hours willing to work for women (simple slope = -2.347, $t(497) = -3.10$, $p = .002$), and a positive and strong interaction ($b = 1.948$, $t(497) = 1.803$, $p = .072$) reversed the trend for men (simple slope = -0.399, $t(497) = -0.518$, $p = .605$). The interaction term was in the same direction and of similar magnitude as the parameter estimate in Study 1.

Table 15: Modeling Hours Willing to Work for Top STEM Doctoral Students

	Dependent variable:				
	Hours in Ideal Job				
	Cohort 1	Cohort 2	Cohort 3	Talent Search	Top STEM Graduate Students
Number of Biological Children	-2.39*** (-3.49, -1.29)	-3.54*** (-5.38, -1.70)	-3.85** (-7.46, -0.25)	-2.80*** (-3.73, -1.87)	-2.35*** (-3.83, -0.86)
Gender	5.73*** (2.68, 8.78)	5.40** (1.079, 9.72)	1.40 (-6.14, 8.94)	5.26*** (2.88, 7.64)	3.14 (-1.03, 7.31)
Interaction	2.34*** (0.91, 3.77)	3.17*** (1.03, 5.31)	2.76 (-1.29, 6.81)	2.48*** (1.34, 3.62)	1.95* (-0.17, 4.06)
Constant	47.19*** (44.85, 49.54)	47.72*** (44.10, 51.34)	52.63*** (46.11, 59.15)	47.91*** (46.00, 49.82)	48.13*** (45.28, 50.98)
Observations	1,081	462	309	1,852	501
Log Likelihood	-4,466.12	-1,856.76	-1,320.73	-7,657.56	-2,036.56

Note. Consistency is shown between Study 1 and Study 2. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

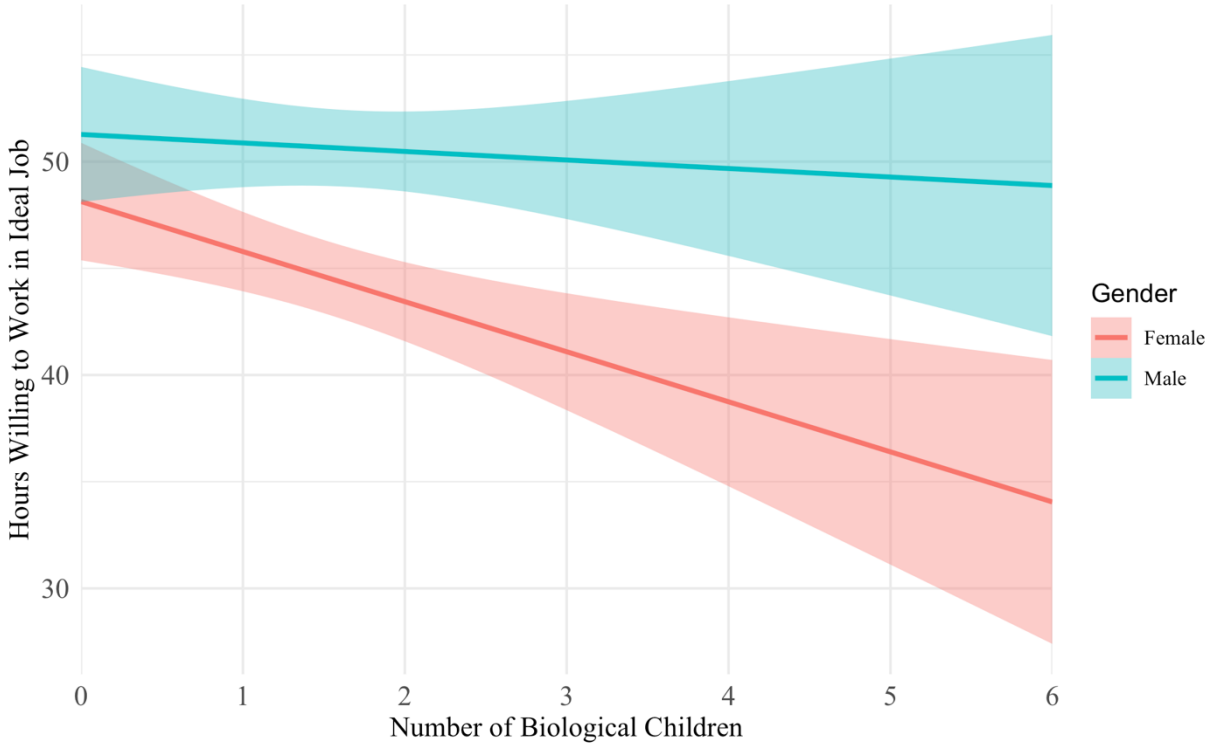

Figure 22. Study 2: Hours Willing to Work as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Subjective Well-Being

As in Study 1, parallel analysis was conducted to derive an ideal linear combination of five indicators of psychological well-being. Figure 23 shows the results. Only the first eigenvalue of the correlation matrix of the psychological well-being measures was above what would be expected by chance. Thus, principal component analysis was conducted and the first component was retained (see Table 16 for a summary of the principal component structure).

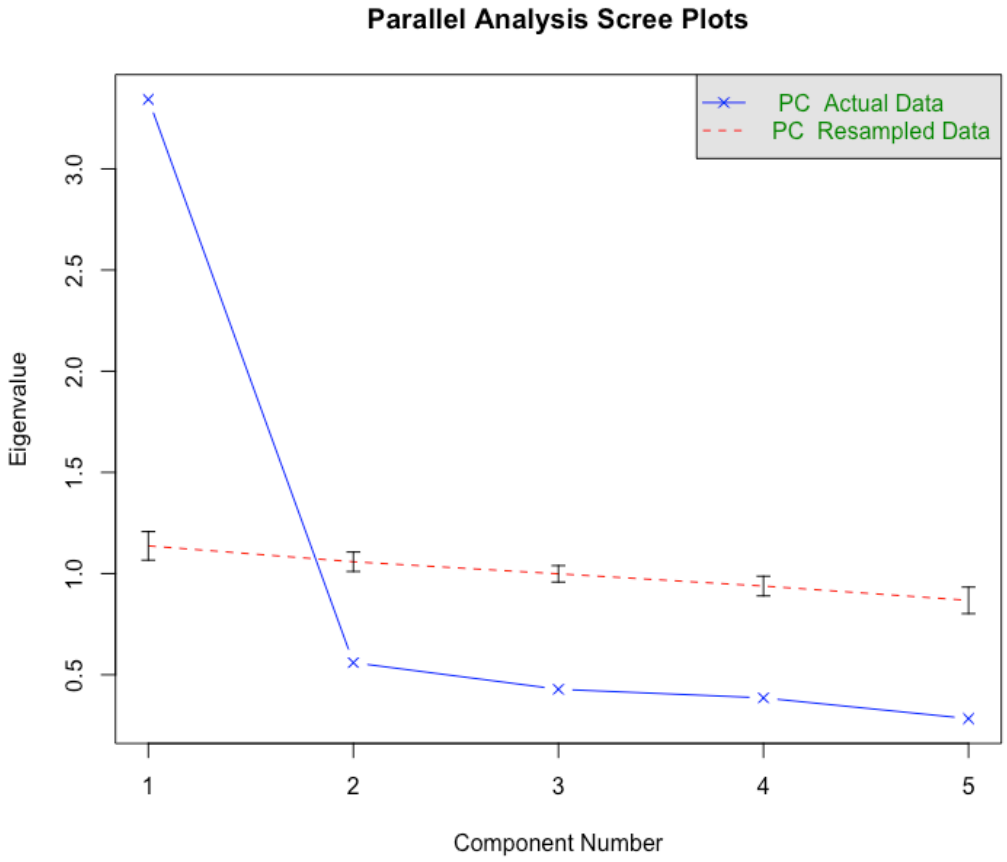

Figure 23. Scree Plot and Parallel Analysis for Top STEM Doctoral Students.
Performed on the correlation matrix of the indicators of psychological well-being for Study 2, only the first eigenvalue is above what would be expected by chance.

Table 16: Principal Component Analysis Results for Top STEM Doctoral Students

	Component Loading	Communality	Uniqueness
Core Self-Evaluations	0.86	0.74	0.26
Positive Affect	0.80	0.63	0.37
Life Satisfaction	0.83	0.69	0.31
Negative Affect (Reversed)	0.80	0.63	0.37
Psychological Flourishing	0.80	0.65	0.35

Note. All indicators had relatively strong and uniform loadings on the first principal component, suggesting relatively equal weighting in its computation.

Table 17 summarizes the regression results of psychological well-being principal component as a function of number of children by gender; Figure 24 shows the relations graphically. Parameter estimates were consistent with Study 1: more children were associated with greater psychological well-being for women (simple slope = 0.169, $t(450) = 2.913$, $p = .004$). As in Study 1, the interaction was virtually zero ($b = -0.053$, $t(450) = -0.664$, $p = .507$), resulting in commensurate increases in psychological well-being as a function of number of children for men (simple slope = 0.116, $t(450) = 2.099$, $p = .036$).

Table 17: Modeling Psychological Well-Being for Top STEM Doctoral Students

	Dependent variable:				
	Psychological Well-Being				
	Cohort 1	Cohort 2	Cohort 3	Talent Search	Top STEM Graduate Students
Number of Biological Children	0.11*** (0.05, 0.18)	0.22*** (0.07, 0.37)	0.13 (-0.10, 0.36)	0.14*** (0.08, 0.20)	0.17*** (0.06, 0.28)
Gender	-0.02 (-0.21, 0.17)	-0.05 (-0.39, 0.30)	-0.08 (-0.57, 0.41)	-0.07 (-0.23, 0.09)	0.20 (-0.11, 0.51)
Interaction	0.03 (-0.06, 0.12)	-0.04 (-0.22, 0.13)	-0.04 (-0.30, 0.23)	0.01 (-0.07, 0.08)	-0.05 (-0.21, 0.10)
Constant	-0.12 (-0.27, 0.02)	-0.29* (-0.58, 0.01)	-0.33 (-0.75, 0.08)	-0.18*** (-0.31, -0.05)	-0.32*** (-0.54, -0.10)
Observations	944	420	284	1,648	454
Log Likelihood	-1,244.39	-599.43	-429.84	-2,295.64	-638.24

Note. Consistency is shown between Study 1 and Study 2. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

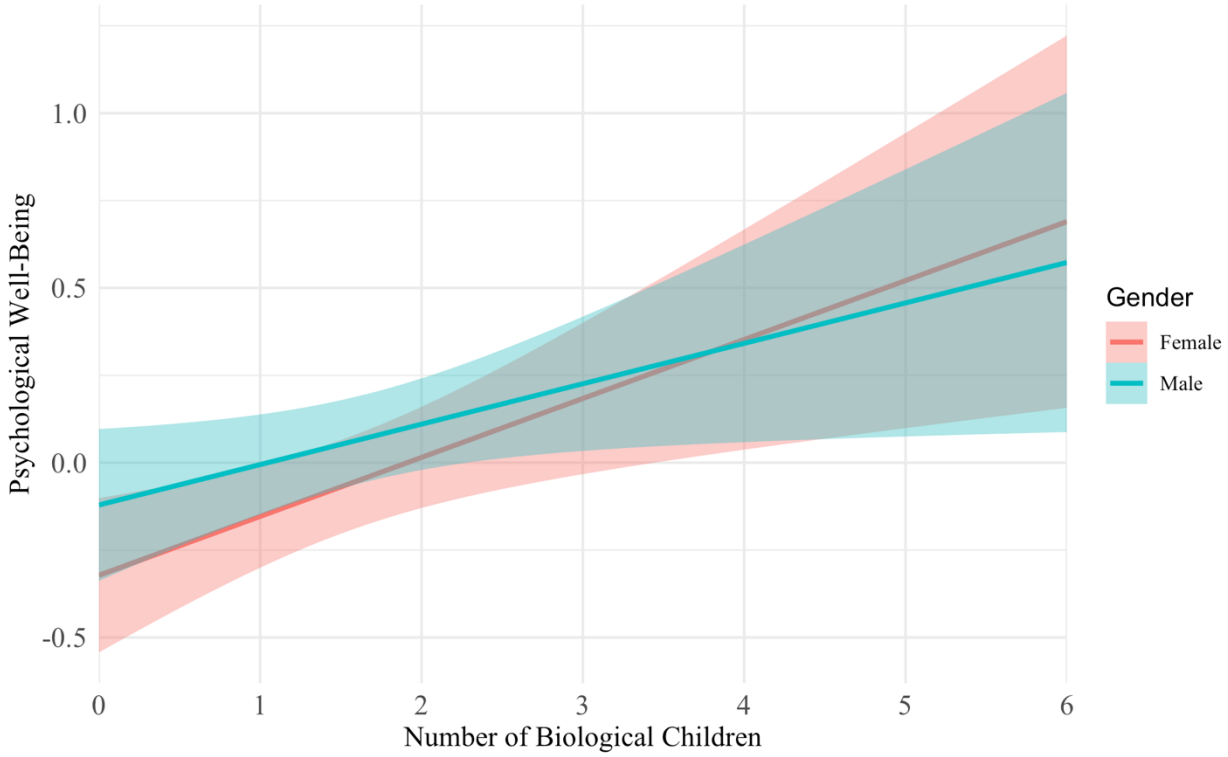

Figure 24. Study 2: Psychological Well-Being as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Figure 25 summarizes mean levels of relationship satisfaction faceted by number of children and gender. As in Study 1, there was very little variation in this one-item measure; most participants described themselves as either satisfied or very satisfied. As a result, it was difficult to test whether men and women with more children were more satisfied in their relationships than those with fewer or no children. Nevertheless, as in Study 1, participants in relationships were generally satisfied.

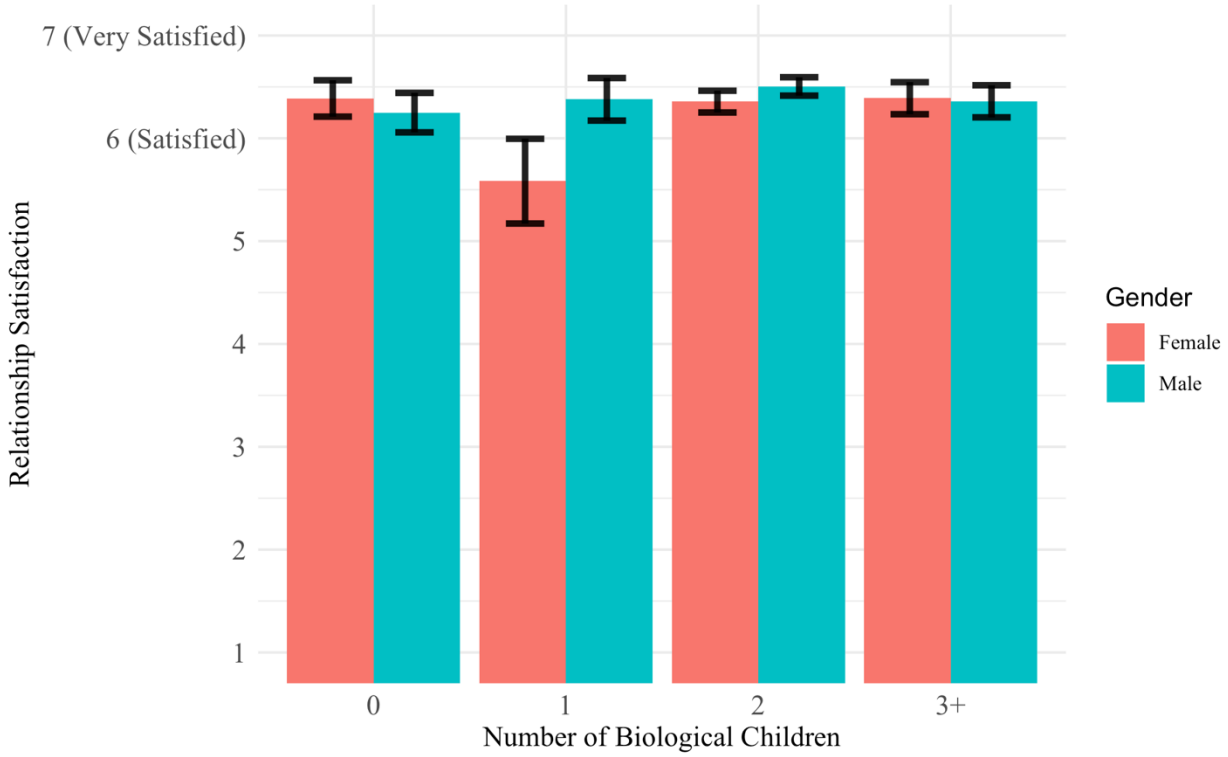

Figure 25. Study 2: Relationship Satisfaction as a Function of Number of Children. Error bars represent ± 1 standard error of the mean.

Table 18 summarizes the regression results of life satisfaction on number of children and gender; Figure 26 shows the relations graphically. Parameter estimates were consistent with Study 1: more children were associated with greater life satisfaction for women (simple slope = 0.928, $t(471) = 2.889$, $p = .004$). As in Study 1, the interaction was virtually zero ($b = -0.228$, $t(471) = -0.510$, $p = .610$), resulting in commensurate increases in life satisfaction as a function of number of children for men (simple slope = 0.699, $t(471) = 2.245$, $p = .025$).

Table 18: Modeling Life Satisfaction for Top STEM Doctoral Students

	Dependent variable:				
	Life Satisfaction				
	Cohort 1	Cohort 2	Cohort 3	Talent Search	Top STEM Graduate Students
Number of Biological Children	1.00*** (0.56, 1.45)	1.32*** (0.40, 2.24)	0.89 (-0.49, 2.28)	1.03*** (0.64, 1.41)	0.93*** (0.30, 1.56)
Gender	-0.43 (-1.69, 0.82)	-0.66 (-2.76, 1.43)	-2.32 (-5.26, 0.61)	-0.81 (-1.81, 0.19)	0.24 (-1.48, 1.96)
Interaction	0.12 (-0.47, 0.71)	-0.28 (-1.35, 0.78)	0.05 (-1.51, 1.61)	0.04 (-0.44, 0.52)	-0.23 (-1.11, 0.65)
Constant	24.25*** (23.30, 25.20)	24.38*** (22.62, 26.15)	25.59*** (23.06, 28.11)	24.48*** (23.68, 25.28)	25.18*** (23.97, 26.39)
Observations	1,018	441	302	1,761	475
Log Likelihood	-3,288.96	-1,433.44	-1,002.90	-5,728.37	-1,495.01

Note. Consistency is shown between Study 1 and Study 2. 95% confidence intervals are given in parentheses. *$p < .10$, **$p < .05$, ***$p < .01$.

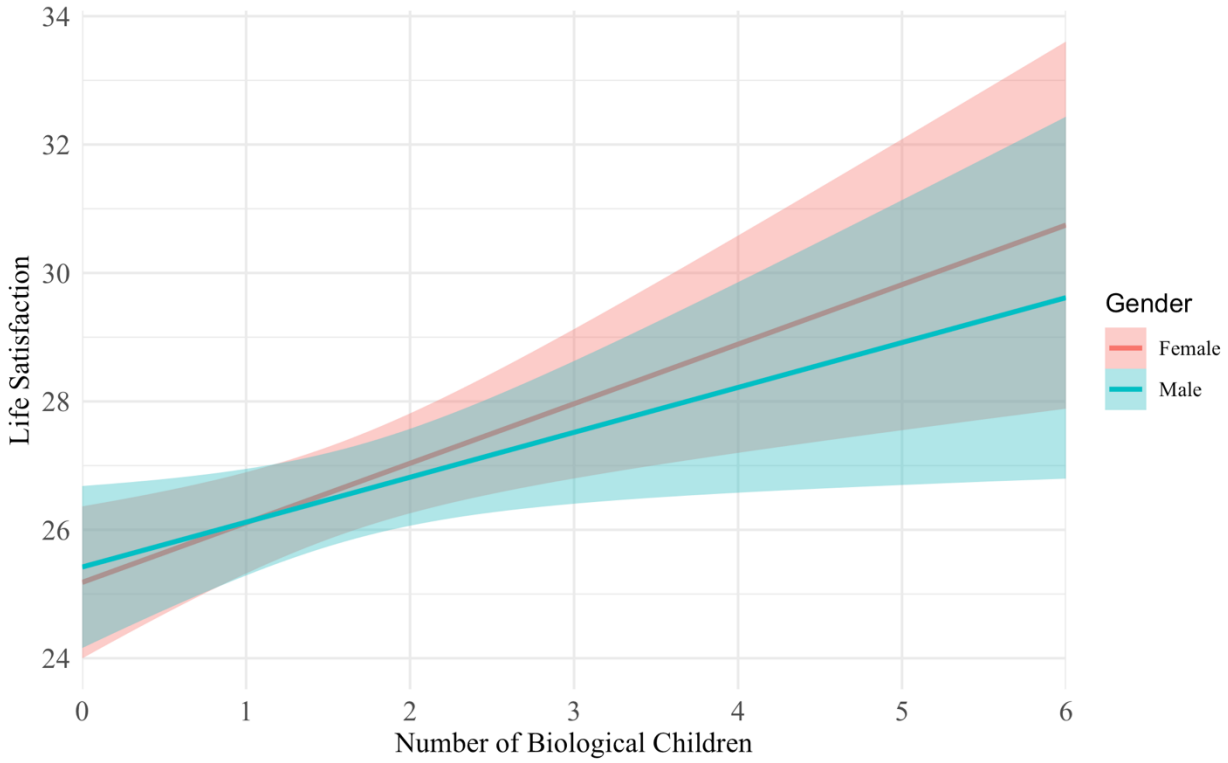

Figure 26. Study 2: Life Satisfaction as a Function of Number of Children. Shaded bands give 95% confidence intervals about the regression line.

Ancillary Analyses

As in Study 1, the female and male top STEM doctoral students in Study 2 saw differential effects on occupational outcomes resulting from having children; nevertheless, both generally saw increases in subjective well-being as a function of having children. As was done in Study 1, several ancillary analyses are conducted here to test whether there is an economic division of labor underlying these seemingly inconsistent findings.

Figure 27A shows, for male and female top STEM doctoral students, that as the number of children for men increases, the probability of having a significant other working full-time decreases; conversely, as women have more children, their significant others are more likely to work full time. Consistent with Study 1 (see Figure 14) this pattern appears to suggest a division

of labor that reconciles the previous findings of divergence in income for mothers and fathers as a function of children coinciding with increases in subjective well-being for both: mothers and fathers simply adapt to changes in demands at home in different ways.

Testing whether this division of labor between partners in a relationship results in decreases in economic well-being, Figure 27B shows that for both male and female top STEM doctoral students, total family income (e.g., participant and spouse income; investment income) does not decrease as participants have more children. Just as in Study 1, these ancillary analyses help to explain the seemingly inconsistent findings of differential effects of children on occupational outcomes for high-potential men and women combined with their similar effects for subjective well-being.

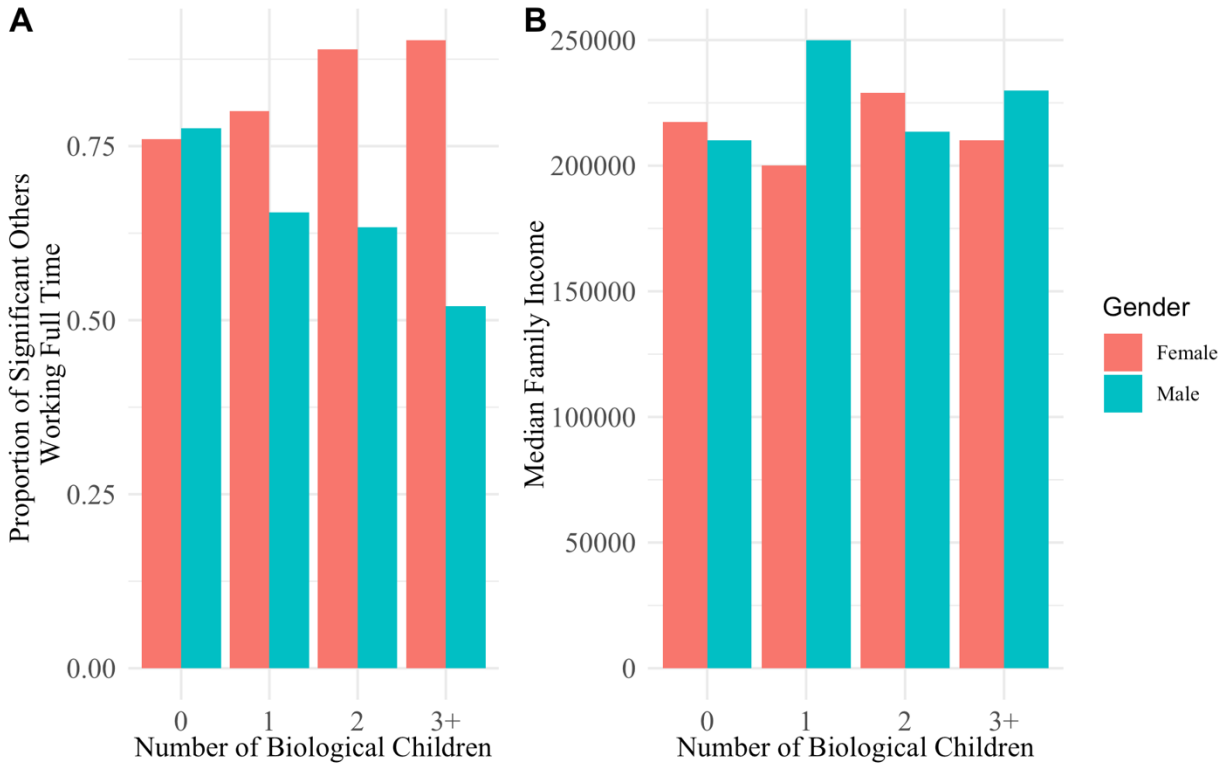

Figure 27. Division of Labor and Family Income amongst Top STEM Doctoral Students
Part A shows diverging trends in the proportion of male and female participants' significant others working full time as a function of number of children of the participant. Despite this division of labor, Part B shows relative constancy in the median family income earned.

Discussion

The results from Study 2 replicate those of Study 1 among elite STEM doctoral students in all important respects. Again, consistent with the literature on the Motherhood Penalty and Fatherhood Bonus (Budig & England, 2001; Cukrowska-Torzewska & Matysiak, 2020), female STEM doctoral students ultimately had substantially lower incomes as a function of the number of children that they had, whereas male STEM doctoral students manifested the inverse trend (earning substantially larger incomes with more children). These findings add robustness to the findings of Study 1 with a distinct high-potential population, and support hypothesis H_{A1} that the incomes of men and women would covary negatively as a function of having children. Income

differs between the genders as a function of parenthood while, simultaneously, attendant outcomes converge across relationship satisfaction, psychological well-being, total family income, and life satisfaction.

Among these elite STEM doctoral students, as income changes diverged between the genders as a function of number of children, both genders experienced overall increases in their psychological well-being and life satisfaction. In addition, their relationship satisfaction remained the same (consistently high) and, even more impressive than the findings of Study 1, the median total family income for STEM doctoral students was above $200,000 across zero to three or more children. By and large, these families are not hurting economically. Like Study 1, as the male participants in Study 2 had more children the likelihood of their partners working full time decreased, whereas the inverse of this relationship was true for female participants. Collectively, however, an impressive total family income was maintained overall. Unlike the trend in Study 1, though, total family income did not covary significantly with children. Total family income was unvaryingly high. In both Study 1 and Study 2, across 1 to 3 or more children, the family median income was over three times that of the average US household.

The prediction of log-odds of leadership and probability of leadership increased for men with more children and remained relatively constant for women with more children; these results were not statistically significant, but they were consistent with the effects found in Study 1. As with the results in Study 1, these findings provide only mixed evidence in support of H_{A2} (which stated that mothers would also see a decrement to likelihood of leadership). Even for these elite STEM doctoral students, it is possible that the criteria for occupational leadership were too stringent and there was too little variability to find individual or group differences as a function of number of children; alternatively, it may be simply true that high-potential mothers'

probabilities of leadership are not substantially decreased as a function of more children (namely, H_{A2} is false).

Study 2's work-life structure results replicated those of Study 1 and provide further support for the second set of hypotheses. Just like intellectually prodigious women, top STEM female doctoral students with more children ultimately reported spending fewer hours at work and more hours with their families and homes; they also reported a willingness to work fewer hours if given their ideal job. Fathers did not work statistically more hours as children came along, did not spend fewer hours at home with the family, and were not less willing to work more hours if given their ideal job. Consistent with Study 1, these findings give statistical support to H_{B1}, H_{B2}, and H_{B3}. That is, high-potential female participants elect to spend more time with family and the home (relative to at work) as more children come along; and high-potential male participants elect to spend more time at work (relative to family and the home) as more children come along.

Finally, Study 2's results on subjective well-being were consistent with those of Study 1. As with the intellectually precocious youth, both male and female participants in Study 2 reported greater psychological well-being and life satisfaction as they had more children (thus providing support for H_{C1} and H_{C3}). However, as in Study 1, the one-item measure of Relationship Satisfaction revealed little variability, and most participants generally scored high on this item; therefore, no conclusions can be made regarding H_{C2} (which stated that both male and female participants would derive greater relationship satisfaction as a function of more children), due to ceiling effects on this indicator. Nevertheless, it is clear from the uniformly high male/female means on this item—whether they had multiple children, one, or none—that both genders were generally happy with their relationships. As in Study 1, and especially when

combined with other findings on high-potential samples in the psychological literature (Ferriman et al., 2009; Gino et al., 2015; Lubinski et al., 2014, under review; Susan Pinker, 2008; Rhoads, 2004), this finding suggests that there are many different ways in which high-potential women and men decide to construct their lives and relationships, which appear to work for them psychologically.

As in Study 1, the ancillary analyses tie together the findings that both high-potential men and women with more children derived greater subjective well-being from having them, while their work-life activities responded in contrasting ways to more children (and the women saw decreases to occupational achievement). The proportion of male participants' significant others working full time decreased as more children came along, and the reverse was true for the significant others of women. Nevertheless, participants and their partners did not earn less as a family as more children came along. Indeed, across the board, their family income was uniformly high (median income greater than $200,000).

A division of labor between participants and their significant others—in which one partner takes on more responsibility in earning income while the other partner takes on more responsibility with the family—seems to be a trend associated with parenthood and, especially, as the number of children in a family grow. Consistent with the results of Study 1, the consequence of this division of labor is not a decrease to the economic resources for the family. Rather, the partnership generally appears to work well on multiple dimensions for high-potential populations and is associated with increases in overall subjective well-being and the maintenance of satisfying romantic relationships.

General Discussion

Parenthood tends to affect the occupational outcomes of high-potential women and men in different ways. The life-changing event of parenthood also affects both women and men in more ways than one. If this series of studies only analyzed gender differences in individual income as a function of parenthood and number of children, knowledge would have been added to the empirical literature on the Motherhood Penalty and Fatherhood Bonus through generalization probes replicated operationally and conceptually across multiple high-potential samples. These two phenomena were firmly replicated among high-potential samples. This is a contribution to the field that also fills the gap on these phenomena for highly educated (Anderson et al., 2002, 2003; Todd, 2001; Wilde et al., 2010) and highly compensated populations (Budig & Hodges, 2010; England et al., 2016; Glauber, 2018; Killewald & Bearak, 2014). Prior research on this topic for high preforming populations has been equivocal, but the findings reported here are consistent and robust.

However, this dissertation goes beyond simply documenting that the Motherhood Penalty and Fatherhood Bonus generalizes to those who are seen as possessing a high degree of human capital. By examining a broad range of other important life outcomes, how this phenomenon operates across high-potential samples can be seen as a more nuanced and richer story. When work-life structure, relationship satisfaction, subjective well-being, division of labor, and total family income are considered jointly, gender differences in individual income for high-potential populations is somewhat less psychologically concerning than is typically portrayed in the social science literature. Indeed, it is conceivable that for some individuals and families, this divergence is associated with more positive benefits than negative deficits overall. By just focusing on

economic considerations, for example, across both studies and all cohorts, gender differences in individual income did not appear to accompany negative financial consequences for total family income. Among the intellectually prodigious cohorts in Study 1, median family income centered around $150,000 for people with children, which was significantly higher than for people with no children. For Study 2, the elite STEM doctoral student's median family income centered around $200,000 regardless of children. When these findings are aligned with how participants felt about their lives and relationships and their self-reports of psychological well-being, a psychological story of interpersonal connectedness, psychological health, and productivity emerges.

Across the samples examined here, the women and men who are earning differential incomes with more children are also feeling better about the lives that they have developed for themselves. To see this more fully requires contextualizing life course phenomena more broadly and taking multiple outcomes into account. Recall the Park et al. (2013) study on educational acceleration that initially focused on one outcome as well: long-term accomplishments and creativity in STEM. When viewed in isolation, males seemed to be profiting more from educational acceleration than females. (This was pausing, because early in their development the girls reported that they enjoyed these accelerative learning experiences as much as the boys—and possibly even more.) But when their occupational outcomes were assessed more broadly, it was found that these women had a greater tendency to specialize in medicine and law and, hence, STEM occupational outcomes used to validate the educational efficacy for acceleration were not inclusive enough to capture their career development and how they chose to invest their talents. When educational and occupational outcome criteria were assessed more comprehensively, the

initial concern was allayed and psychologically understood. This seems to be the case with other life course phenomena as well.

Just as Kahneman and Tversky (Kahneman, 2011; Kahneman et al., 1982; Kahneman & Tversky, 2000) have shown for economic decision making, people do not always behave to maximize profit because other considerations are frequently at play. This too may be the case for the dynamic properties of family units. Yet, even so, across both studies and all cohorts examined here, the observed gender differences in individual income did not appear to result in negative economic consequences for family income. As a function of more children, male participants had significant others who tended to work less, while female participants tended to have significant others who worked more; and in both situations, total family income did not appear to decrease as a function of more children. So, not only was assessing additional outcomes informative, but assessing income at a different level of molarity with ancillary analyses was clarifying as well. As other research on high-potential populations has shown (Gino et al., 2015; Lubinski et al., 2014, under review; Susan Pinker, 2008), expanding psychological appraisals of outcome criteria, and aggregating assessments for multiple levels of analysis, reveals that understanding lives fully requires broad assessments at multiple levels of analysis. Unidimensional assessments are unlikely to ever tell the full story.

The importance of a broad and holistic view of decision-making is particularly informative in high-potential populations because these populations have the most options available to them in terms of educational, career, and life choices. Putting too narrow of a scope on the ways in which these individuals could define success (e.g., by only considering their *individual* income) would give the inaccurate psychological portrait that many participants somehow failed in their lives and were unable to find fulfillment. Even within the realm of

economic outcomes, however, a broader perspective, like moving from individual- to family income, affords a better understanding of the lives that people live. Such considerations are fertile soil for future research because many individuals make concessions to their individual salaries for other economic and personal reasons such as working remotely versus in the office for efficiency, being close to family, living in a certain region of the country, or spending more time with family and less time at work. Other scenarios may be played out as well.

Consider, for example, the many full-time careers that are characterized by non-linear pay scales. In many careers, an X percent increase in hours worked leads to a larger-than-X percent increase in the amount of income brought home. Many business and legal professions are characterized by these pay structures, with many employees seeing substantial increases in pay for putting in extra hours on evenings or being willing to work on call over weekends. By contrast, a linear pay scale career is one in which an X% increase in hours devoted to the career is compensated with approximately an X% increase in income (Goldin, 2014).

Non-linear pay scale careers tend to be precisely the types of careers which are largely populated by high-potential individuals such as those examined here. Considering participants and their significant others as a team and, given a certain amount of life tasks that need to be accomplished in a given week (e.g., earning income, raising children, and caring for the home and extended family), it is not surprising to find that high-potential participants and their partners might elect for a differential division of labor between the home and work (recall Figure 14A and Figure 27A), when it still allows for a lucrative total family income overall (recall Figure 14B and Figure 27B). This could happen, especially for partners on non-linear pay scales, if one partner increases their number of hours at work from, say, 40 to 60 hours per week in order to double their income whereas the other partner on a linear pay scale decreases their work from 40

hours per week to 20, 10, or 0. Indeed, partners could conceivably earn the same amount of total family income with fewer combined hours spent working per week, or in many cases earn more total family income! An example of this phenomenon was detailed by Miller (2014) in *The New York Times* and is likely a realistic scenario in many households considered in this dissertation. There are other considerations and realistic scenarios as well.

For example, the cost of childcare in the United States is exceptionally high, and the benefits of being closer to family members in order to help with caring for children can be economically and emotionally rewarding. Many people elect to live closer to family members for childcare or more optimal family considerations in terms of contact with loved ones, even if these decisions come at the cost of decrements to salary. A relatively small decrement in salary or occupational prestige could be markedly offset by such highly valued family resources. Often discussed lifestyle changes during the COVID-19 pandemic makes balancing these life personal/professional utilities an especially intriguing research topic. Family income, like individual income, is not a perfect indicator of all tangible economic resources or emotional resources for a family's overall well-being. It does not, for example, capture the emotional support and advantages of having loving grandparents nearby. At the very least, this underscores the importance of contextualizing a wide range of life course considerations and outcomes.

Another reason this study is a particularly informative is that it is the longest-running longitudinal study of the lifespan development of high-potential women who largely grew up during and after second-wave feminism. Relative to Lewis Terman's landmark "Genetic Studies of Genius" (Holahan et al., 1995; Terman, 1926; Terman & Oden, 1959), the women in this study had not only the ability to excel in a variety of different live endeavors, but also the *opportunity* to do so. Indeed, the women in Study 2 were trained in some of the best STEM

graduate training institutions in the world. With increases in opportunity come increases in the number of ways in which individuality can be expressed (Dawis, 1992; Tyler, 1992; Williamson, 1965); and, like divergences in educational credentials (Lubinski, 2020; Stoet & Geary, 2018, 2020), lifestyle divergences were indeed observed.

Of course, the women in Studies 1 and 2 faced challenges that the men did not, and they were not promised an easy road to success in life. That bias in the workplace still exists is undeniable. Yet, to suggest that they were tricked into accepting a lower salary and a worse quality of life seems questionable, considering their own self-reports of psychological well-being, life satisfaction, and relationship satisfaction here and elsewhere (Ferriman et al., 2009; Lubinski et al., 2014, in press), and coupled with the total family income that they enjoyed.

Psychological Findings and Policy Implications

The findings uncovered in Studies 1 and 2 have important psychological implications for the way in which the Gender Wage Gap is conceptualized and interrupted. In terms of studying life course developmental phenomena and understanding differential outcomes from a psychological point of view, a broader perspective is informative for understanding psychological well-being and financial health as well as relationship satisfaction and family dynamics. Policy formation is different, however; it is based on values and what society hopes to achieve. Scientific findings can inform how best to achieve our goals, but they cannot tell us which goals we should value (Humphreys, 1995). Some of the literature on the Gender Wage Gap conflates these two matters, but they are important to uncouple for analytic and other purposes.

For example, the standard narrative surrounding the Gender Wage Gap today is that the gap is necessarily a societal ill and evidence of injustice and discrimination against women (El-Hout et al., 2021). In fact, substantial efforts and attention are currently being allocated to eliminating the Gender Wage Gap (e.g., Equal Pay Day, Paycheck Fairness Act). However, the series of findings in this dissertation suggest that a more guarded interpretation of the Gender Wage Gap and Motherhood Penalty is in order, which does not dictate policy but certainly has policy relevance. Indeed, in some proposals for reducing the gender wage gap, it is difficult to discern whether what is being recommended is for individuals to change their educational, occupational, or lifestyle preferences (and "lean in") or whether something else is being recommended. The particulars are often unclear (El-Hout et al., 2021).

Others have suggested that more comprehensive treatment of gender differences be conducted for the same reasons. Thus recall, at the conclusion of his leadership role working on the U.S. National Academies Report, "Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty" (2010), co-chair Claude Canizares commented in an interview published in *Science* on the committee's empirical findings: "While women can take some encouragement from the fact that there is no evidence of large-scale bias at these key transition points, the reasons for their continued underrepresentation need to be examined more closely" (Mervis, 2009, p. 1250). Canizares hoped that federal agencies and universities would gather longitudinal data on the career paths of women and concluded his remarks by saying, "I'd suggest we start with our own graduate students" (Mervis, 2009, p. 1251).

The current study does just that for the most academically and scientifically accomplished STEM doctoral students of their generation and three cohorts of individuals

originally identified as profoundly gifted 12-year-olds. What it shows is that to have a more complete understanding of gender differences in a host of occupational outcomes, determinants well beyond what is required for understanding educational/occupational choice and compensation need to be taken into account. This is something that economic theorists and evolutionary psychologists have recommended for decades for conceptualizing gender-differentiating trends (Browne, 2002; Buss, 2019; Geary, 2021; Hakim, 2017; Murray, 2020; Rhoads, 2004). The current findings highlight the need to consider these broader theoretical points of view.

Finally, it is important to distinguish between a general description of what *is* the state of affairs for high-potential women and men in this study and what *ought* to be the state of affairs. This distinction is Hume's *is-ought problem* in which statements about what is the case do not necessarily determine what ought to be the case, which necessarily involves value judgments and morals (Black, 1964). The emphasis of this dissertation is in describing what is currently the case for high-potential mothers and fathers, how differences and similarities in important life outcomes come about and inform psychological understandings. Prescriptions of what ought to be the case necessarily bring into account values and ethics. Science can inform best practices for how to reach valued social goals, but it does not dictate what ought to be done. It is quite apparent that these two sets of considerations are highly blended in contemporary literature, and care must be taken to keep them separate. Consider a point of view that Erich Fromm (1956) shared in *The Art of Loving*:

> The intellectual has one prime task to fulfill, first, last, and always. It is his job to search out the truth as best he can and to speak that truth. It is not the intellectual's primary calling, it is not his primary function, to draft political platforms… But it is the intellectual's special task—and this is what defines his role or should define it—to pursue the truth without compromise and without regard for his own or anyone else's interests. If intellectuals restrict their function of finding and speaking the whole truth in the service of any program or any political goals, no matter how praiseworthy the program or the goals may be, then those intellectuals are failing in their own unique task and, ultimately, in the most important political task they have. For I feel that political progress depends on how much of the truth we know, how clearly and boldly we speak it, and how great an impression it makes on other people. (p. 116)

Replication and Consistency of Findings

A key strength to this dissertation is the replication features that were internalized in its design. Especially now amidst the replication crisis in psychology (Camerer et al., 2018), it is important for investigators to consider consistency and stability in not only statistical parameter estimates, but also in relations among higher-order constructs. It has been well documented that many research findings in psychology specifically and the soft sciences in general simply do not replicate (Ioannidis, 2005). This makes it increasingly difficult to interpret research findings in the psychological sciences with confidence. Without a series of replicated results and consistency in these results, findings should be interpreted cautiously. Consistency in findings is precisely what the series of operational replications buttressed by a constructive replication were designed to reveal.

Study 1 demonstrated a consistent pattern of findings across three separate cohorts of increasingly more able youth followed from age 13 until age 50. General findings emerged in the Study that, as in less select populations, intellectually prodigious women with more children saw decreases in financial remuneration, decreases in hours worked and willing to work, and increases in hours spent with the family. However, across these three cohorts, these same women

also seem to derive more subjective well-being from their life decisions than did their peers with fewer or no children. A very different series of results for men consistently came out of the analyses across the three cohorts. Men with more children tended to actually have more favorable occupational outcomes; and they had relatively few changes in hours worked, hours willing to work in their ideal job, and hours spent with family. Nevertheless, men with more children—just like women with more children—tended to report more subjective well-being.

Thus, the results in Study 1 successfully constituted a series of operational replications according to Lykken's (1968, 1991) three-tiered framework of replications with increasing degrees of scientific credence (literal → operational → constructive). That is, three cohorts of intellectually precocious youth, identified in similar ways and yet over different time points and degrees of intellectual selectivity, gave rise to consistent patterns of gender differences and similarities as a function of having children.

Ultimately, given the consistency in estimates amongst these three cohorts of intellectually precocious youth, the cohorts were combined into a general sample of intellectually precocious youth and regression parameter estimates were estimated based on this combined sample for the most accurate and precise inferences. These parameter estimates are the best available from a sample of nearly 2,000 intellectually precocious youth, and it is these estimates that should be referenced in future investigations of the constructs of interest in this dissertation. Indeed, this is precisely what Study 2 had ultimately been designed to accomplish.

The replication conducted by Study 2 is similar in spirit to those conducted in Study 1, but the findings from Study 2 are ultimately significantly more compelling when taken in combination with the findings in Study 1. These participants had not only much more psychological similarity across genders, but they also were highly uniform in their remarkable

and STEM concentrated educational histories. That consistent empirical patterns were uncovered across Study 1 and Study 2 reinforces an underlying pattern of multivariate relationships of the constructs structuring the lifespan development of high-potential men and women. That is, the results of this study suggest these patterns found in high-potential populations (e.g., intellectually precocious, high STEM potential) dovetail with the general population.

Study 2 demonstrated, in a sample of highly select top STEM graduate students, results that were consistent with those in Study 1. As with the intellectually prodigious women, female top STEM graduate students with more children saw decreases in financial remuneration, decreases in hours worked and willing to work, and increases in hours spent with the family. Nevertheless, as in Study 1, these women derived more subjective well-being from their life decisions than did those women with fewer or no children. The effect of children on men was markedly different. Those men with more children earned robustly larger incomes but had relatively few changes in hours worked, hours willing to work in their ideal job, and hours spent with family. Nevertheless, these male top STEM graduate students with more children, just like their female peers in Study 2, reported more subjective well-being regarding the way their lives had developed.

The consistency of findings within Study 1 and the consistency of findings across Studies 1 and 2 combined to suggest a robust and general pattern of relationships that can inform future theorizing about important societal questions regarding the Gender Wage Gap, Motherhood Penalty, Fatherhood Bonus, and other important life outcomes in high-potential populations in general.

Limitations and Future Directions

Future investigators would do well to address some limitations in the current research. The most obvious limitation to this dissertation is the lack of ability to make firm causal claims regarding the directional effects. In each analysis conducted, several plausible causal mechanisms could be considered. For example, two plausible interpretations of the finding that women with more biological children earn less are 1. That these women decided to have more children as a result of being less engaged with their careers and 2. That these women became disengaged from their careers after beginning to have more children. Clearly, both mechanisms could be operating simultaneously to differing degrees.

To complicate matters further, even if the causal direction amongst the components of this relationship were determined, a totally different causal mechanism could be conceivable for men. For example, it seems plausible from an evolutionary psychology perspective that men who achieve greater career success also succeed more often in having children. It is also possible that the presence of children per se leads to greater feelings of responsibilities for fathers and hence greater career success (i.e., a mediational mechanism between children and remuneration). The same complexities of interpretation obviously extend to the findings for subjective well-being (do those who are happier go on to have more children, or do the children per se lead to greater happiness among high-potential individuals).

Future studies of this population should attempt to make use of longitudinal data, when available, to make stronger inferences regarding the direction of effects amongst the constructs at play. For example, one could use survival analysis to compare whether and when participants in this sample had children as a function of relatively early indicators of career promise and

remuneration. Additionally, one could study longitudinal changes in subjective well-being as a function of when participants had children.

The scope of this dissertation is both a strength and a limitation of the design of this dissertation. On the one hand, this dissertation has allowed for the first holistic investigation of a multivariate array of important life outcomes in high-potential samples in the context of the Gender Wage Gap. This allows for clear and focused inferences on high-potential populations specifically, and these populations have rarely been given serious attention in investigations of the Gender Wage Gap and Motherhood Penalty.

Nevertheless, care must be given to not extrapolate the pattern of multivariate findings into samples on which the inferences were not made. Holistic approaches for broader populations are still lacking, and these findings do not necessarily inform theorizing about the relationship of remuneration, familial commitment, and subjective well-being in the broader population. Future researchers interested in holistically addressing the determinants of the Gender Wage Gap and the ways in which the Gender Wage Gap interacts with other important outcomes should do so in a more broadly representative sample (using, for example, nationally representative data) in order to determine the extent to which these findings generalize beyond the specific population of interest in this dissertation.

Finally, future investigations should consider additional control variables which could be used in multivariate investigations of the phenomena at play. Much is now known about the determinants of gender differences in pay including ability, interests, sector of employment, presence of children, and discrimination. Although many of these components were controlled implicitly in the design of this study (for example, by using samples selected on ability or interest), more complete controls could be completed by explicitly including measures of these

important determinants in future models of wages regressed on key constructs of interest (but see also Bernstein and colleagues, 2021 and Meehl, 1970 for a discussion of issues arising in the artificial or statistical control of such variables).

Conclusion

The Gender Wage Gap has consistently emerged as one of the most important questions we face as a modern society (Giles, 2011, p. 18); this phenomenon is particularly important to explore in high-potential populations, from which many men and women will come to make many of the most important contributions in an increasingly conceptual economy (Friedman, 2005; Hunt, 1995, 2011; Zakaria, 2008). Moreover, gender differences in high-power leadership positions are most conspicuous at the top of multiple disciplines and professions. By exploring an important covariate of occupational differences between men and women (parenthood and children), while simultaneously assessing some of the most important concomitant outcomes in structuring a life (work-life structure and subjective well-being), this dissertation has taken a holistic approach to conceptualizing the many ways in which high-potential populations structure their lives and how they feel about their decisions.

There are many different ways of creating and pursuing a meaningful life, and decisions regarding career and family are likely the two most conspicuous arenas in this process. Discrimination in the workplace should be decried, but the Gender Wage Gap per se should not be taken as prima facie evidence for discrimination. Instead, researchers and the public in general need to consider the diversity of ways in which people decide to structure their lives and, most importantly, how they feel about the decisions they have made. Appreciation for human individuality and diversity in decision making necessitates holistic evaluations of outcomes and the rationales behind them.

APPENDIX A.

Leverage and Influence of Observations Across Models

Figure 28 through Figure 34 show the leverage and influence of observations across the regression models conducted in this dissertation. Following the diagnostic strategy of Hoaglin and Welsch (1978), Panels A and B of each figure show the leverage and influence, respectively, of individual observations in Study 1; Panels C and D show the same for Study 2. Note that, because the model matrix **X** contains only fourteen unique rows (that is, the fourteen combinations of 0-6 biological children and two genders), and because the diagonal elements of $\mathbf{H} = \mathbf{X}(\mathbf{X}^{T}\mathbf{X})^{-1}\mathbf{X}^{T}$ (i.e., individual leverage values) are only a function of the model matrix **X**, only fourteen unique leverage values are possible for each of Panels A and C. Naturally, observations with more biological children had larger leverage values across these analyses; nevertheless, the influence of these observations was not generally any greater than the influence of observations at lower numbers of children, nor were their standardized residuals typically found outside ±2. Thus, in order to conduct inferences at relatively large numbers of children (e.g., 4, 5, and 6)—and without extrapolating uncomfortably into sparsely supported ranges of the data (e.g., 7, 8, 9, 10, and 11 children)—I do not collapse any observations with six or fewer children.

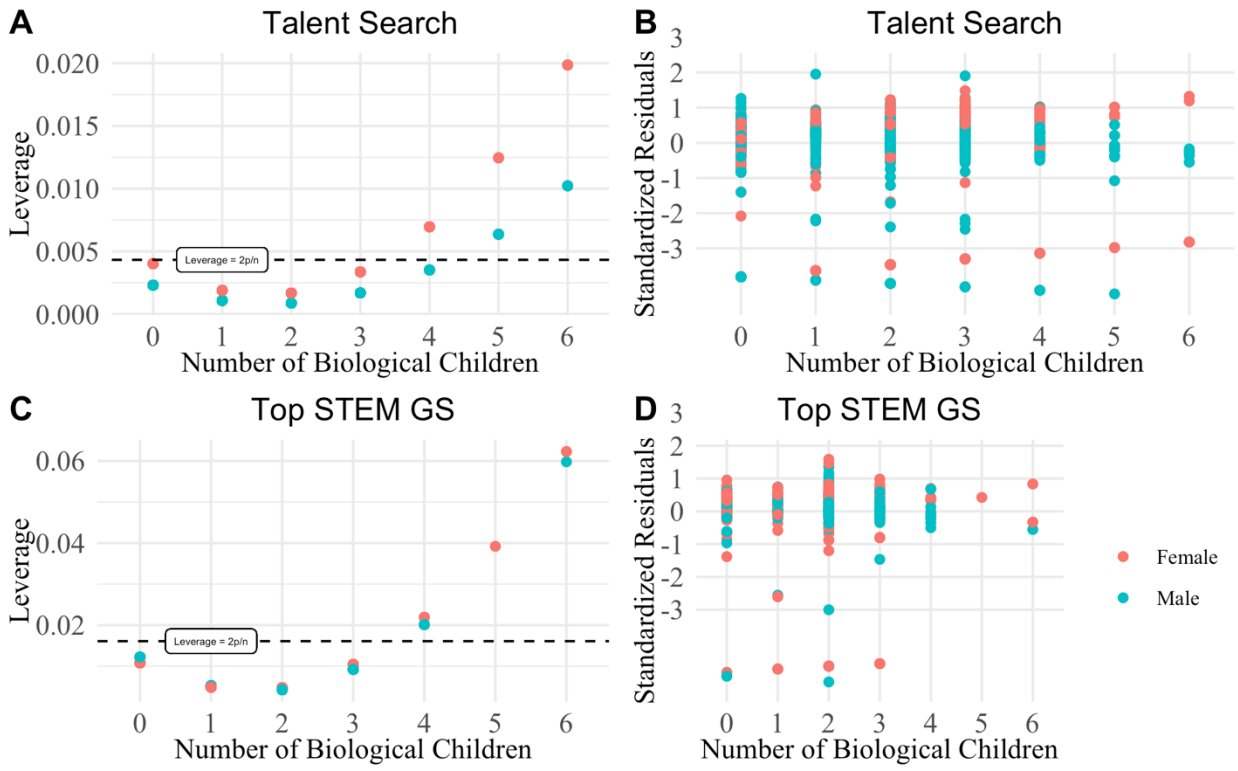

Figure 28. Leverage and Influence for Income Analysis.

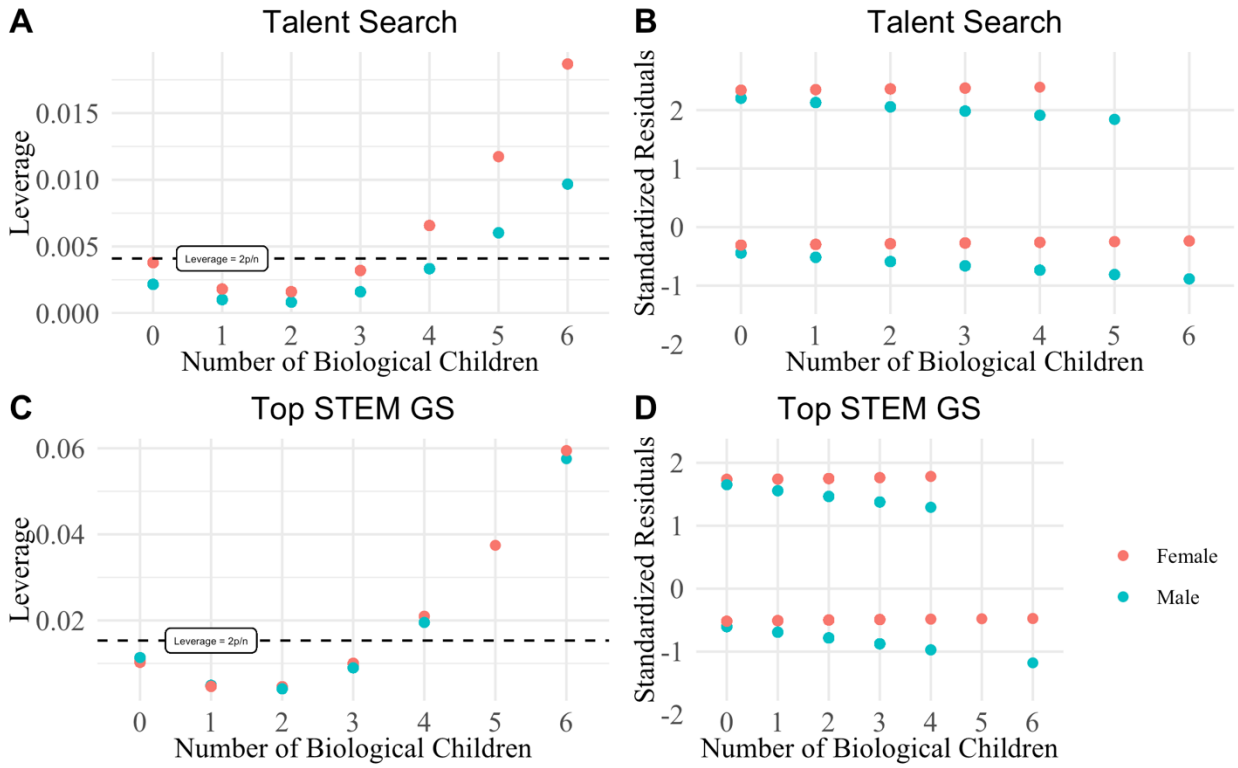

Figure 29. Leverage and Influence for Leadership Analysis.

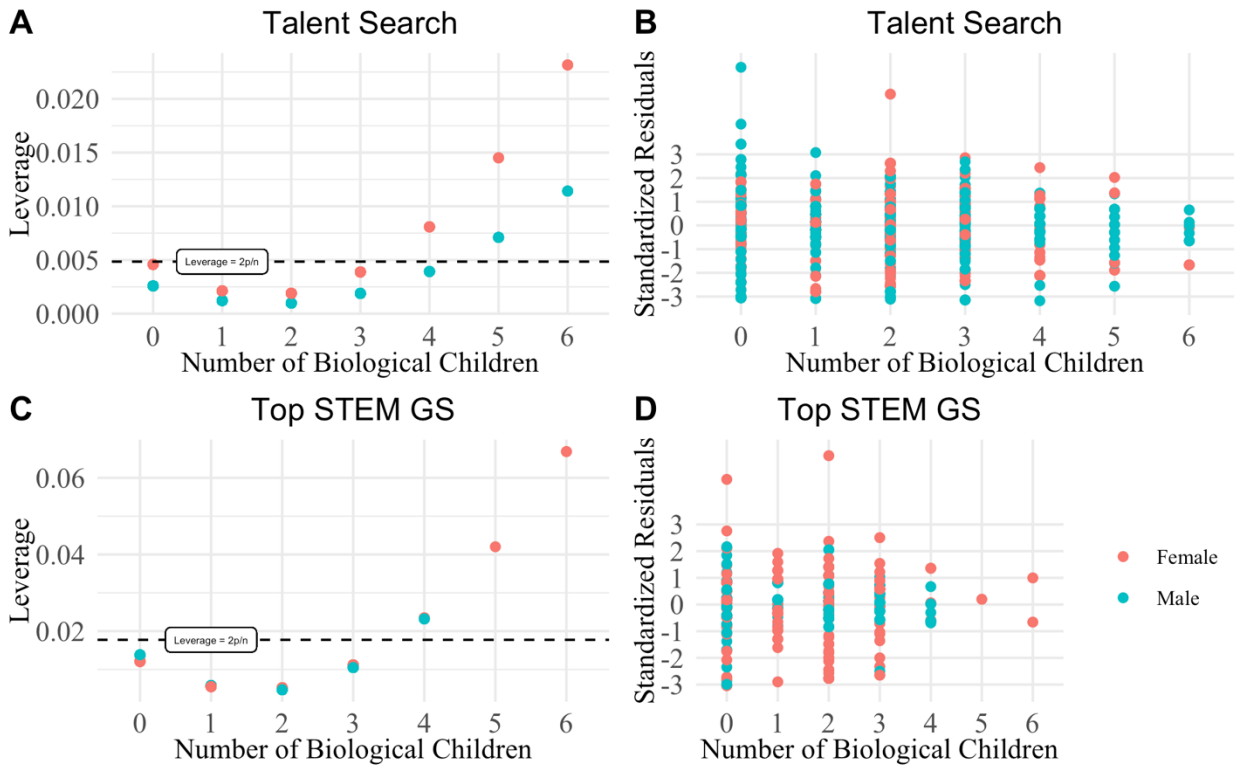

Figure 30. Leverage and Influence for Hours Worked Analysis.

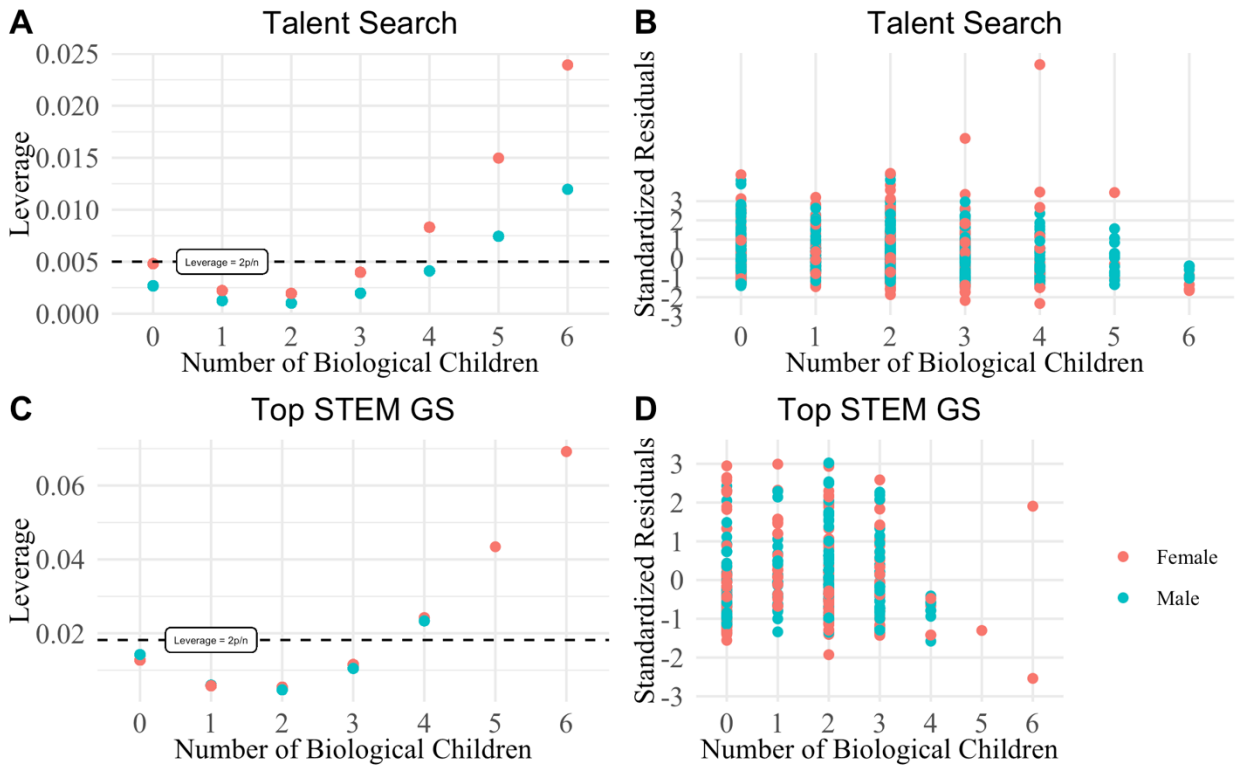

Figure 31. Leverage and Influence for Hours with Family and Home Analysis.

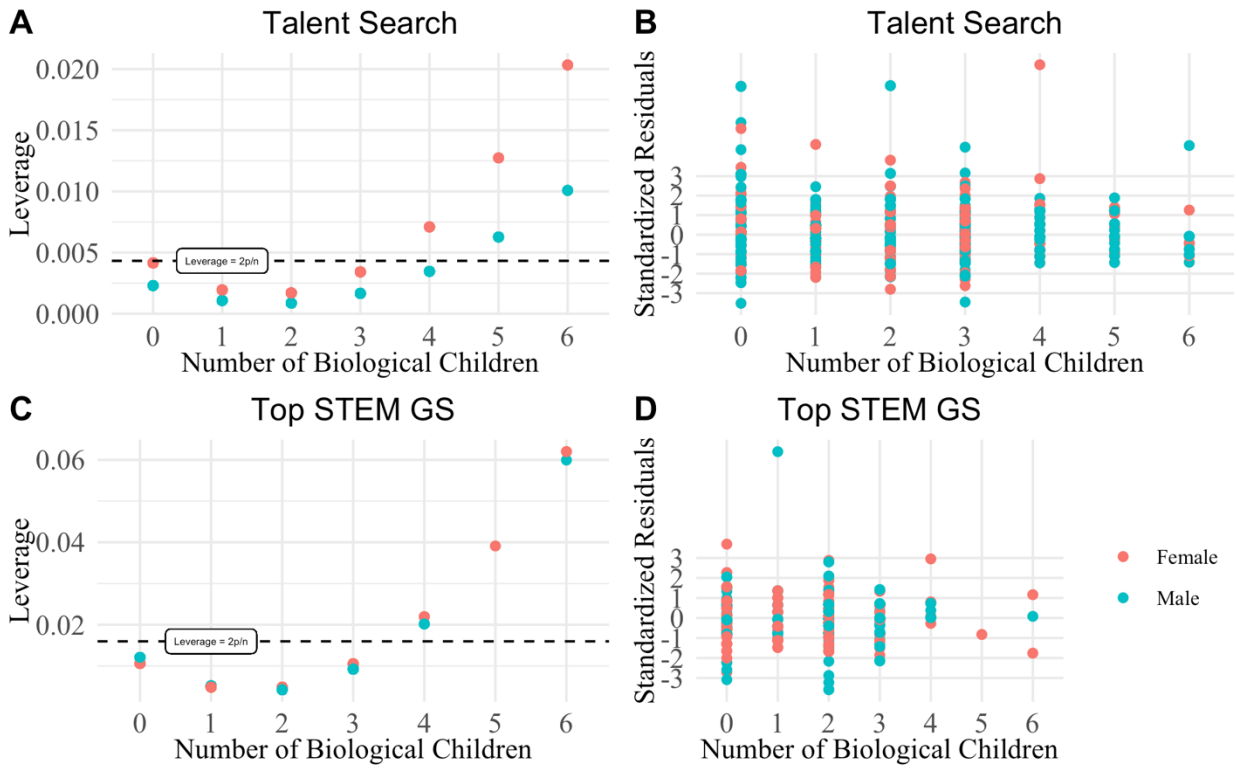

Figure 32. Leverage and Influence for Hours Willing to Work in Ideal Job Analysis.

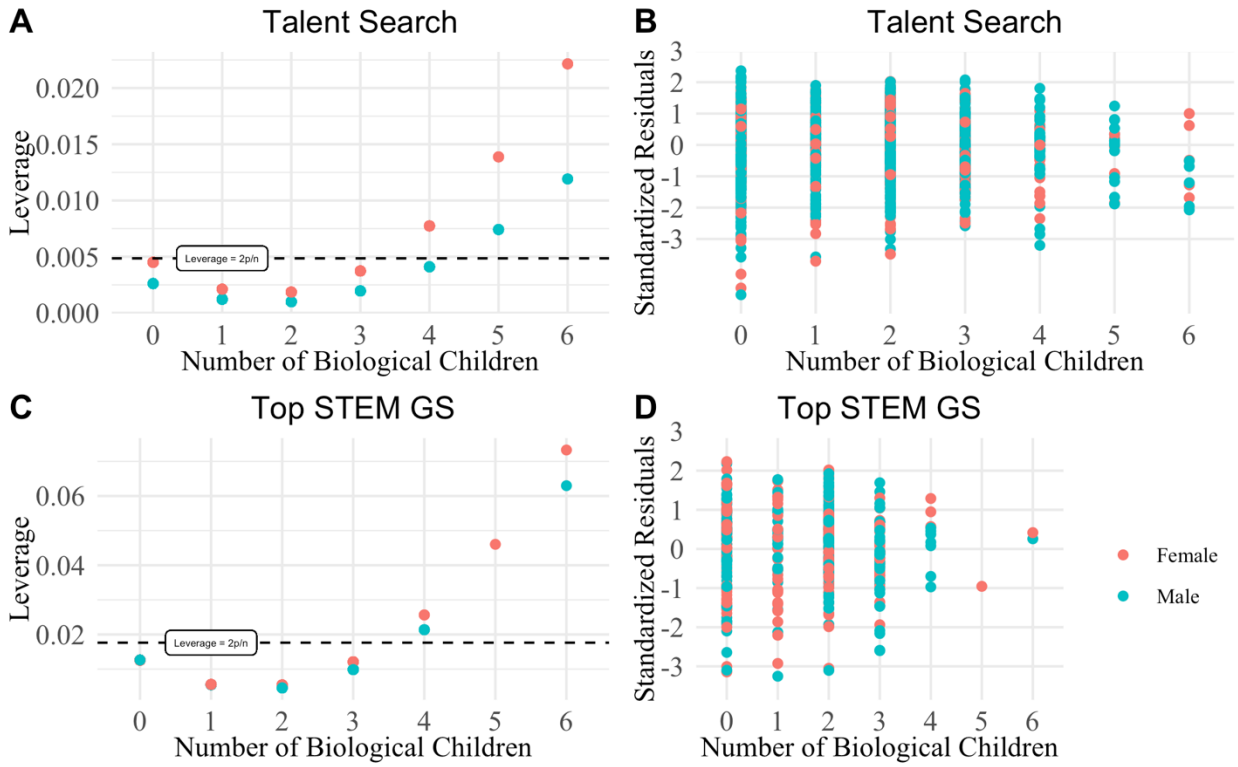

Figure 33. Leverage and Influence for Psychological Well-Being Analysis.

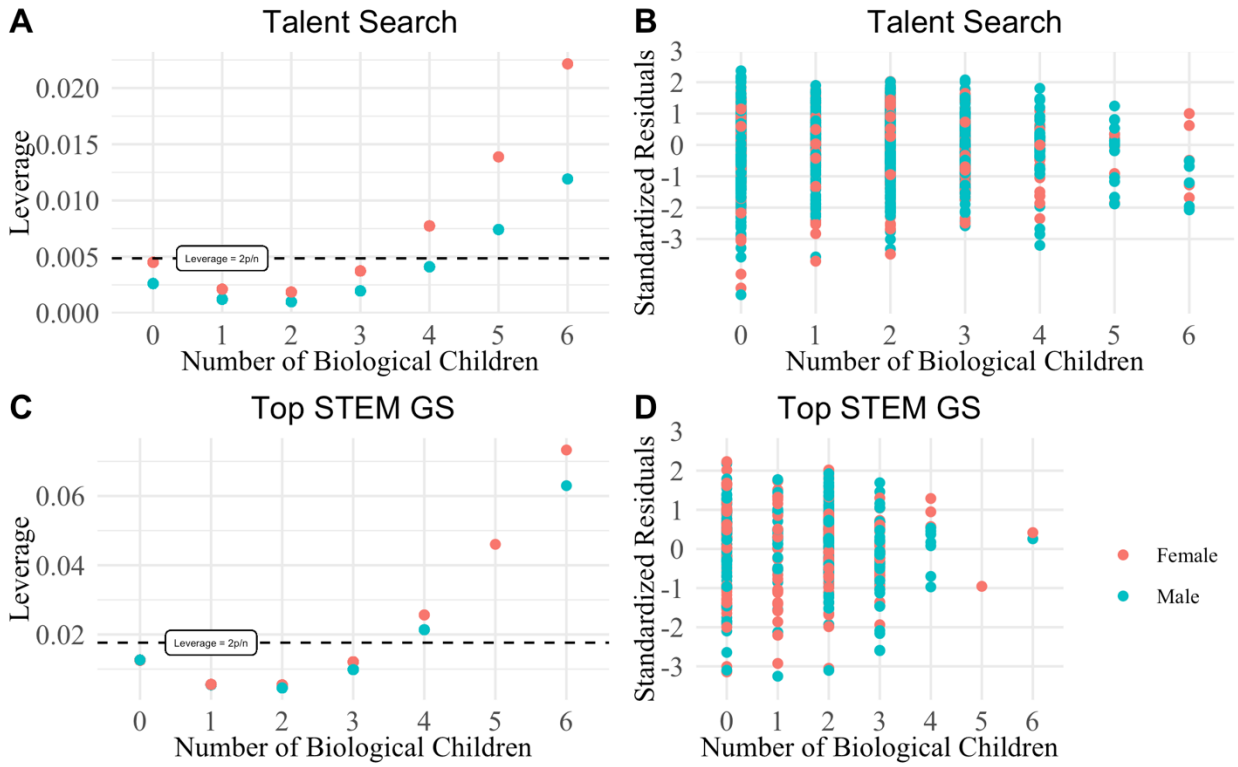

Figure 34. Leverage and Influence for Life Satisfaction Analysis.

APPENDIX B.

Alternative Functional Forms

It is natural to ask whether the relationship between income and number of biological children for mothers and fathers might be non-linear. While it seems highly likely that the relationship is monotonic (e.g., increases in children lead to consistent decreases in income for mothers), it seems at first improbable that the relationship is linear (e.g., the change from four children to five has the same impact on income as does the change from zero children to one). Exploring non-linearity could come in the form of Box-Cox transformations (Box & Cox, 1964; e.g., taking the logarithm of wages). Moreover, an investigator might consider several alternatives for modeling such as qualitative effects of motherhood, polynomial regression, and generalized additive models (GAM; Ruppert et al., 2003).

Investigators have also warned in particular about non-linear trends masquerading as interactions when modeled in a linear fashion (Busemeyer & Jones, 1983; Cortina, 1993; Lubinski & Humphreys, 1990). This is of particular concern when two interactants are highly correlated (thereby leading to high correlations between their product term and their individual quadratic terms).

Following the recommendation of Gelman and colleagues in looking for non-linear trends (2021, pp. 161-162), Figure 35 shows residuals plotted against predicted values in the regression of log-wages on number of biological children, gender, and their interaction for Study

1 and Study 2. In both studies, the residuals seem to be randomly dispersed around the mean (zero), homoscedastic, and uncorrelated with the predicted values—thus providing support for the assumptions regarding errors in linear regression. Moreover, the point-biserial correlation between number of biological children and gender (male coded 1) was $r(1950) = .031$, $p = .172$ for Study 1 and $r(520) = .051$, $p = .247$ for Study 2; these low values dispel the concerns by previous researchers (Busemeyer & Jones, 1983; Cortina, 1993; Lubinski & Humphreys, 1990) regarding the likelihood of non-linear effects disguised as interactions. These results, taken in combination with an effort to integrate my findings into the literature on the Motherhood Penalty (which consistently models log wages as a linear function of children), suggest that a linear functional form is appropriate in this dissertation (see, for example, Cukrowska-Torzewska and Matysiak, 2020 for meta-analysis or Budig and England, 2001 for a discussion of this issue).

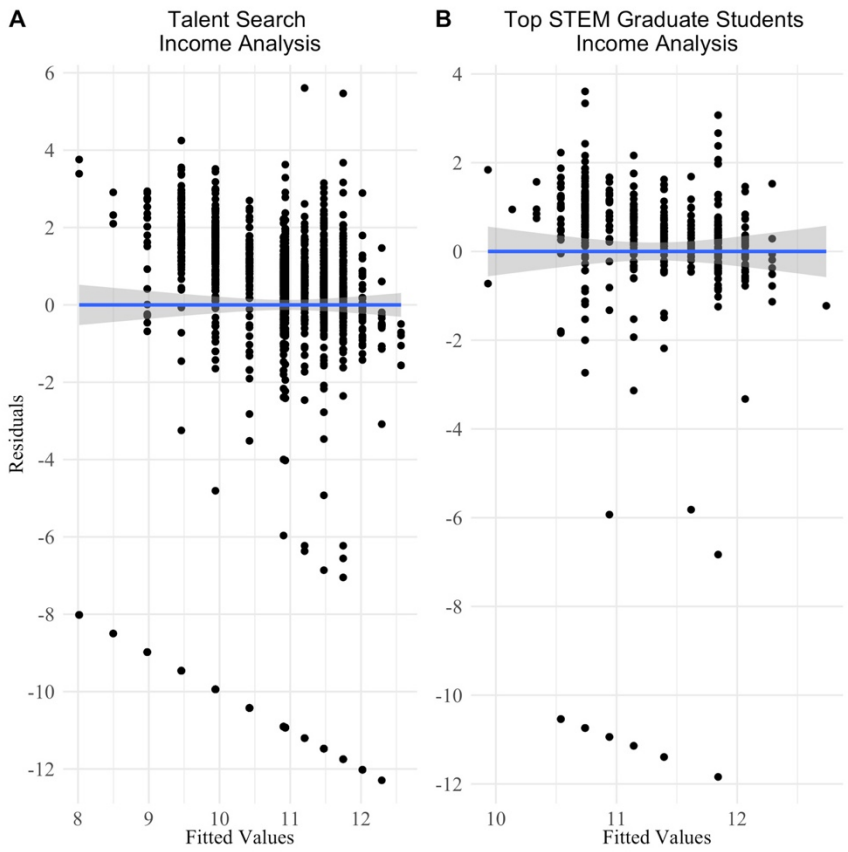

Figure 35. Looking for Evidence of Non-Linearity in Income Analyses.

REFERENCES

Abelson, R. P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin*, *97*, 129–133. https://doi.org/10.1037/0033-2909.97.1.129

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*, 123–129. https://doi.org/10.1016/S0165-1765(03)00032-6

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage Publications.

Anderson, D. J., Binder, M., & Krause, K. (2002). The motherhood wage penalty: Which mothers pay it and why? *The American Economic Review*, *92*, 354–358. https://doi.org/10.1257/000282802320191606

Anderson, D. J., Binder, M., & Krause, K. (2003). The motherhood wage penalty revisited: Experience, heterogeneity, work effort, and work-schedule flexibility. *Industrial & Labor Relations Review*, *56*, 273–294. https://doi.org/10.2307/3590938

Arden, R., & Plomin, R. (2006). Sex differences in variance of intelligence across childhood. *Personality and Individual Differences*, *41*, 39–48. https://doi.org/10.1016/j.paid.2005.11.027

Azmat, G., & Ferrer, R. (2017). Gender gaps in performance: Evidence from young lawyers. *The Journal of Political Economy*, *125*, 1306–1355. https://doi.org/10.1086/693686

Becker, E., & Lindsay, C. M. (2004). Assortative mating or glass ceiling: Under-representation of female workers among top earners. In S. W. Polachek (Ed.), *Accounting for Worker Well-Being* (Vol. 23, pp. 235–267). Emerald Group Publishing Limited.

https://doi.org/10.1016/S0147-9121(04)23007-2

Becker, G. S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics*, *3*, S33–S58. https://doi.org/10.1086/298075

Becker, G. S. (1991). *A treatise on the family*. Harvard University Press.

Benbow, C. P., Lubinski, D., Shea, D. L., & Eftekhari-Sanjani, H. (2000). Sex differences in mathematical reasoning ability: Their status 20 years later. *Psychological Science*, *11*, 474–480. https://doi.org/10.1111/1467-9280.00291

Benbow, C. P., & Stanley, J. C. (1996). Inequity in equity: How current educational equity policies place able students at risk. *Psychology, Public Policy, and Law*, *2*, 249–293.

Bernstein, B. O., Lubinski, D., & Benbow, C. P. (2019). Psychological constellations assessed at age 13 predict distinct forms of eminence 35 years later. *Psychological Science*, *30*, 444–454. https://doi.org/10.1177/0956797618822524

Bernstein, B. O., Lubinski, D., & Benbow, C. P. (2021). Academic acceleration in gifted youth and fruitless concerns regarding psychological well-being: A 35-year longitudinal study. *Journal of Educational Psychology*, *113*, 830–845. https://doi.org/10.1037/edu0000500

Bertrand, M., Goldin, C., & Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, *2*, 228–255. https://doi.org/10.1257/app.2.3.228

Black, M. (1964). The gap between "is" and "should." *The Philosophical Review*, *73*, 165–181. https://doi.org/10.2307/2183334

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*, 789–865. https://doi.org/10.1257/jel.20160995

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal*

Statistical Society. Series B (Methodological), *26*, 211–252. http://www.jstor.org/stable/2984418

Brockmann, H., Koch, A.-M., Diederich, A., & Edling, C. (2018). Why managerial women are less happy than managerial men. *Journal of Happiness Studies*, *19*, 755–779. https://doi.org/10.1007/s10902-016-9832-z

Browne, K. R. (2002). *Biology at work: Rethinking sexual equality*. Rutgers University Press.

Browne, K. R. (2005). Women in science: Biological factors should not be ignored. *Cardozo Women's Law Journal*, *11*, 509–528.

Browne, K. R. (2018). The quixotic quest for "gender equality" in the workplace. *University of Toledo Law Review*, *49*, 685–714.

Budig, M. J., & England, P. (2001). The wage penalty for motherhood. *American Sociological Review*, *66*, 204–225. https://doi.org/10.2307/2657415

Budig, M. J., & Hodges, M. J. (2010). Differences in disadvantage: Variation in the motherhood penalty across white women's earnings distribution. *American Sociological Review*, *75*, 705–728. https://doi.org/10.1177/0003122410381593

Bureau of Labor Statistics. (2019). *Women had higher median earnings than men in relatively few occupations in 2018*. The Economics Daily. https://www.bls.gov/opub/ted/2019/women-had-higher-median-earnings-than-men-in-relatively-few-occupations-in-2018.htm

Bureau of Labor Statistics. (2020). *Median weekly earnings of full-time wage and salary workers by detailed occupation and sex*. https://www.bls.gov/cps/cpsaat39.htm

Busemeyer, J. R., & Jones, L. E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, *93*, 549–562.

https://doi.org/10.1037/0033-2909.93.3.549

Buss, D. M. (2019). *Evolutionary psychology: The new science of the mind* (6th ed.). Routledge.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*, 367–383. https://doi.org/10.1037/0033-2909.125.3.367

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., & Altmeid, A. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behavior*, *2*, 637–644. https://doi.org/10.1038/s41562-018-0399-z

Cattell, R. B. (1957). *Personality and motivation structure and measurement*. World Book.

Cech, E. A., & Blair-Loy, M. (2019). The changing career trajectories of new parents in STEM. *Proceedings of the National Academy of Sciences*, *116*, 4182–4187. https://doi.org/10.1073/pnas.1810862116

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, *15*, 75–141. https://doi.org/10.1177/1529100614541236

Ceci, S. J., Kahn, S., & Williams, W. M. (2021). Stewart-Williams and Halsey argue persuasively that gender bias is just one of many causes of women's underrepresentation in science. *European Journal of Personality*, *35*, 40–44. https://doi.org/10.1177/0890207020976778

Ceci, S. J., & Williams, W. M. (2015). Women have substantial advantage in STEM faculty hiring, except when competing against more-accomplished men. *Frontiers in Psychology*, *6*, 1532. https://doi.org/10.3389/fpsyg.2015.01532

Chamberlain, A., Zhao, D., & Stansell, A. (2019). *Progress on the gender pay gap: 2019*. https://www.glassdoor.com/research/app/uploads/sites/2/2019/03/Gender-Pay-Gap-2019-Research-Report-1.pdf

Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, *83*, 50–58. https://doi.org/10.1016/j.jebo.2011.06.007

Cheryan, S., & Markus, H. R. (2020). Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychological Review*, *127*, 1022–1052. https://doi.org/10.1037/rev0000209

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, *143*, 1–35. https://doi.org/10.1037/bul0000052

Clynes, T. (2016). How to raise a genius: A long-running study of exceptional children reveals what it takes to produce the scientists who will lead the twenty-first century. *Nature*, *573*, 152–155.

Cook, C., Diamond, R., Hall, J., List, J. A., & Oyer, P. (2018). *The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers* (No. 24732; National Bureau of Economic Research). https://doi.org/10.3386/w24732

Corbett, C., & Hill, C. (2012). *Graduating to a pay gap: The earnings of women and men one year after college graduation*.

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *The American Journal of Sociology*, *112*, 1297–1338. https://doi.org/10.1086/511799

Cortina, J. M. (1993). Interaction, nonlinearity, and multicollinearity: Implications for multiple regression. *Journal of Management*, *19*, 915–922.

https://doi.org/10.1177/014920639301900411

Cukrowska-Torzewska, E., & Matysiak, A. (2020). The motherhood wage penalty: A meta-analysis. *Social Science Research*, *88–89*, 102416. https://doi.org/10.1016/j.ssresearch.2020.102416

Dawis, R. (1992). The individual differences tradition in counseling psychology. *Journal of Counseling Psychology*, *39*, 7–19. https://doi.org/10.1037/0022-0167.39.1.7

Dawis, R. V., & Lofquist, L. H. (1984). *A psychological theory of work adjustment*. University of Minnesota Press.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, *49*, 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. won, Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, *97*, 143–156. https://doi.org/10.1007/s11205-009-9493-y

El-Hout, M., Garr-Schultz, A., & Cheryan, S. (2021). Beyond biology: The importance of cultural factors in explaining gender disparities in STEM preferences. *European Journal of Personality*, *35*, 45–50. https://doi.org/10.1177/0890207020980934

England, P., Bearak, J., Budig, M. J., & Hodges, M. J. (2016). Do highly paid, highly skilled women experience the largest motherhood penalty? *American Sociological Review*, *81*, 1161–1189. https://doi.org/10.1177/0003122416673598

Ferriman, K., Lubinski, D., & Benbow, C. P. (2009). Work preferences, life values, and personal views of top math/science graduate students and the profoundly gifted: Developmental

changes and gender differences during emerging adulthood and parenthood. *Journal of Personality and Social Psychology*, *97*, 517–532. https://doi.org/10.1037/a0016030

Friedman, T. L. (2005). *The World is Flat: A Brief History of The Twenty-First Century*. Farrar, Straus, and Giroux.

Fromm, E. (1956). *The art of loving*. Harper Perennial.

Geary, D. C. (2021). *Male, female: The evolution of human sex differences* (3rd ed.). American Psychological Association.

Gelman, A., Hill, J., & Vehtari, A. (2021). *Regression and other stories*. Cambridge University Press.

Giles, J. (2011). Social science lines up its biggest challenges: "Top ten" crucial questions set research priorities for the field. *Nature*, *470*, 18–19. https://doi.org/doi:10.1038/470018a

Gino, F., Wilmuth, C. A., & Brooks, A. W. (2015). Compared to men, women view professional advancement as equally attainable, but less desirable. *Proceedings of the National Academy of Sciences*, *112*, 12354–12359. https://doi.org/10.1073/pnas.1502567112

Glauber, R. (2018). Trends in the motherhood wage penalty and fatherhood wage premium for low, middle, and high earners. *Demography*, *55*, 1663–1680. https://doi.org/10.1007/s13524-018-0712-5

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, *4*, 26–42. https://doi.org/10.1037/1040-3590.4.1.26

Goldin, C. (2014). A grand gender convergence: Its last chapter. *The American Economic Review*, *104*, 1091–1119. https://doi.org/10.1257/aer.104.4.1091

Goldin, C., & Katz, L. F. (2016). A most egalitarian profession: Pharmacy and the evolution of a family-friendly occupation. *Journal of Labor Economics*, *34*, 705–746.

https://doi.org/10.1086/685505

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *The American Economic Review*, *90*, 715–741. https://doi.org/10.1257/aer.90.4.715

Gough, M., & Noonan, M. (2013). A review of the motherhood wage penalty in the United States. *Sociology Compass*, *7*, 328–342. https://doi.org/10.1111/soc4.12031

Gourman, J. (1989). *The Gourman report: A rating of undergraduate programs in American and international universities*. National Education Standards.

Hakim, C. (2017). *Key issues in women's work: Female diversity and the polarisation of women's employment* (2nd ed.). Routledge-Cavendish.

Harari, Y. N. (2018). *21 lessons for the 21st century*. Spiegel & Grau.

Hardin, J., & Hilbe, J. (2018). *Generalized linear models and extensions (4th ed.)*. STATA Press.

Harzing, A. W. (2007). *Publish or perish*. https://harzing.com/resources/publish-or-perish

Hedges, L. V, & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*, 41–45. https://doi.org/10.1126/science.7604277

Hegewisch, A., & Williams-Baron, E. (2018). *The gender wage gap: 2017 Earnings differences by gender, race, and ethnicity*.

Hoaglin, D. C., & Welsch, R. E. (1978). The hat matrix in regression and ANOVA. *The American Statistician*, *32*, 17–22. https://doi.org/10.1080/00031305.1978.10479237

Holahan, C. K., Sears, R. R., & Cronbach, L. J. (1995). *The gifted group in later maturity*. Stanford University Press.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis.

Psychometrika, *30*, 179–185. https://doi.org/10.1007/BF02289447

Humphreys, L. G. (1988). Sex differences in variability may be more important than sex differences in means. *Behavioral and Brain Sciences*, *11*, 195–196.

Humphreys, L. G. (1995). Foreward. In D. Lubinski & R. V. Dawis (Eds.), *Assessing individual differences in human behavior: New methods, concepts, and findings* (pp. ix–xi). Davis-Black.

Hunt, E. B. (1995). *Will we be smart enough?: A cognitive analysis of the coming workforce.* (1st ed.). Russell Sage Foundation.

Hunt, E. B. (2011). *Human intelligence*. Cambridge University Press.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*, 0696–0701. https://doi.org/10.1371/journal.pmed.0020124

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger.

Johnson, W., Carothers, A., & Deary, I. J. (2008). Sex differences in variability in general intelligence: A new look at the old question. *Perspectives on Psychological Science*, *3*, 518–531. https://doi.org/10.1111/j.1745-6924.2008.00096.x

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The Core Self-Evaluations Scale: Development of a measure. *Personnel Psychology*, *56*, 303–331. https://doi.org/10.1111/j.1744-6570.2003.tb00152.x

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Kahneman, D., & Tversky, A. (2000). *Choices, Values and Frames*. Cambridge University Press.

Kerr, S. P., Kerr, W. R., & Xu, T. (2017). *Personality traits of entrepreneurs: A review of recent literature* (No. 24097; NBER Working Paper Series). https://www.nber.org/papers/w24097

Killewald, A., & Bearak, J. (2014). Is the motherhood penalty larger for low-wage women? A comment on quantile regression. *American Sociological Review*, *79*, 350–357. https://doi.org/10.1177/0003122414524574

Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, *11*, 181–209. https://doi.org/10.1257/app.20180010

Kuncel, N. R., & Hezlett, S. A. (2010). Fact and fiction in cognitive ability testing for admissions and hiring Decisions. *Current Directions in Psychological Science*, *19*, 339–345. https://doi.org/10.1177/0963721410389459

Lofquist, L. H., & Dawis, R. V. (1991). *Essentials of person-environment correspondence counseling*. University of Minnesota Press.

Lubinski, D. (2010). Neglected aspects and truncated appraisals in vocational counseling: Interpreting the interest–efficacy association from a broader perspective: Comment on Armstrong and Vogel (2009). *Journal of Counseling Psychology*, *57*, 226–238. https://doi.org/10.1037/a0019163

Lubinski, D. (2016). From Terman to today: A century of findings on intellectual precocity. *Review of Educational Research*, *86*, 900–944. https://doi.org/10.3102/0034654316675476

Lubinski, D. (2020). Understanding educational, occupational, and creative outcomes requires assessing intra-individual differences in abilities and interests. *Proceedings of the National Academy of Sciences*, *117*, 16720–16722.

Lubinski, D., & Benbow, C. P. (2006). Study of Mathematically Precocious Youth after 35

years: Uncovering antecedents for the development of math-science expertise. *Perspectives on Psychological Science*, *1*, 316–345. https://doi.org/10.1111/j.1745-6916.2006.00019.x

Lubinski, D., & Benbow, C. P. (2021). Intellectual precocity: What have we learned since Terman? *Gifted Child Quarterly*, *65*, 3–28. https://doi.org/10.1177/0016986220925447

Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science*, *25*, 2217–2232. https://doi.org/10.1177/0956797614551371

Lubinski, D., Benbow, C. P., McCabe, K. O., Bernstein, B. O. (under review) Composing meaningful lives: Exceptional women and men at age 50.

Lubinski, D., Benbow, C. P., Shea, D. L., Eftekhari-Sanjani, H., & Halvorson, M. B. J. (2001). Men and women at promise for scientific excellence: Similarity not dissimilarity. *Psychological Science*, *12*, 309–317. https://doi.org/10.1111/1467-9280.00357

Lubinski, D., & Humphreys, L. G. (1990). Assessing spurious "moderator effects": Illustrated substantively with the hypothesized ("synergistic") relation between spatial and mathematical ability. *Psychological Bulletin*, *107*, 385–393. https://doi.org/10.1037//0033-2909.107.3.385

Lubinski, D., & Humphreys, L. G. (1996). Seeing the forest from the trees: When predicting the behavior or status of groups, correlate means. *Psychology, Public Policy, and Law*, *2*, 363–376. https://doi.org/10.1037/1076-8971.2.2.363

Lucas, R. E., Diener, E., & Suh, E. M. (1996). Discriminant validity of well-being measures. *Journal of Personality and Social Psychology*, *71*, 616–628. https://doi.org/10.1037/0022-3514.71.3.616

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin*,

70, 151–159. https://doi.org/10.1037/h0026141

Lykken, D. T. (1991). What's wrong with psychology anyway. In D. Cicchetti & W. M. Grove (Eds.), *Thinking clearly about psychology* (pp. 3–39). University of Minnesota Press.

Mason, M. A., Wolfinger, N. H., & Goulden, M. (2013). *Do babies matter? Gender and family in the ivory tower*. Rutgers University Press.

McCabe, C. J. (2021). *modglm: Computing interaction effects for nonlinear probability and counts*.

McCabe, C. J., Halvorson, M. A., King, K. M., Cao, X., & Kim, D. S. (2021). Interpreting interaction effects in generalized linear models of nonlinear probabilities and counts. *Multivariate Behavioral Research*, 1–27. https://doi.org/10.1080/00273171.2020.1868966

McCabe, K. O., Lubinski, D., & Benbow, C. P. (2020). Who shines most among the brightest?: A 25-year longitudinal study of elite STEM graduate students. *Journal of Personality and Social Psychology*, *119*, 390–416. https://doi.org/http://dx.doi.org/10.1037/pspp0000239

McDowell, J. M. (1982). Obsolescence of knowledge and career publication profiles: Some evidence of differences among fields in costs of interrupted careers. *The American Economic Review*, *72*, 752–768.

Meehl, P. E. (1970). Nuisance variables and the ex post facto design. In M. Radner & S. Winokur (Eds.), *Minnesota studies in the philosophy of science* (Vol. 4, pp. 373–402). University of Minnesota Press.

Mervis, J. (2009). Report finds no gender bias in faculty hiring, resources. *Science*, *324*, 1250–1251. https://doi.org/10.1126/science.324_1250a

Miller, C. C. (2014, September 6). The motherhood penalty vs. the fatherhood bonus. *The New York Times*. https://www.nytimes.com/2014/09/07/upshot/a-child-helps-your-career-if-

youre-a-man.html

Mize, T. (2019). Best practices for estimating, interpreting, and presenting nonlinear interaction effects. *Sociological Science*, *6*, 81–117. https://doi.org/10.15195/v6.a4

Murray, C. A. (2003). *Human accomplishment: The pursuit of excellence in the arts and sciences, 800 B.C. to 1950.* Harper Collins.

Murray, C. A. (2020). *Human diversity: The biology of gender, race, and class.* Twelve.

National Research Council. (1987). *Summary report 1986: Doctoral recipients from United States universities*. National Academy Press.

National Research Council. (2010). *Gender differences at critical transitions in the careers of science, engineering and mathematics faculty*. The National Academies Press.

National Science Foundation. (2021). *What we do*. https://www.nsf.gov/about/what.jsp

Neumark, D., Bank, R. J., & Nort, K. D. Van. (1996). Sex discrimination in restaurant hiring: An audit study. *The Quarterly Journal of Economics*, *111*, 915–941. https://doi.org/10.2307/2946676

Noonan, M. C., Corcoran, M. E., & Courant, P. N. (2005). Pay differences among the highly trained: Cohort differences in the sex gap in lawyers' earnings. *Social Forces*, *84*, 853–872. https://doi.org/10.1353/sof.2006.0021

O'Dea, R. E., Lagisz, M., Jennions, M. D., & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nature Communications*, *9*, 3777. https://doi.org/10.1038/s41467-018-06292-0

Okahana, H., & Zhou, E. (2018). *Graduate enrollment and degrees: 2007 to 2017*. https://cgsnet.org/publication-pdf/5464/CGS_GED17_Report.pdf

Park, G., Lubinski, D., & Benbow, C. P. (2013). When less is more: Effects of grade skipping on

adult STEM productivity among mathematically precocious adolescents. *Journal of Educational Psychology*, *105*, 176–198. https://doi.org/10.1037/a0029481

Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychological Assessment*, *5*, 164–172. https://doi.org/10.1037/1040-3590.5.2.164

Pinker, Steven. (2002). *The blank slate: The modern denial of human nature*. Penguin Putnam.

Pinker, Susan. (2008). *The sexual paradox: Men, women, and the real gender gap*. Scribner.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*, 437–448. https://doi.org/10.3102/10769986031004437

Reich, R. (1991). *The work of nations: Preparing ourselves for the 21st century capitalism*. Knopf.

Revelle, W. (2021). *psych: Procedures for psychological, psychometric, and personality research* (2.1.3). Northwestern University. https://cran.r-project.org/package=psych

Rhoads, S. E. (2004). *Taking sex differences seriously*. Encounter Books.

Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Cambridge University Press.

Ryff, C. D., Singer, B. H., & Love, G. D. (2004). Positive health: connecting well-being with biology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*, 1383–1394. https://doi.org/10.1098/rstb.2004.1521

Sackett, P. R., Lievens, F., Van Iddekinge, C. H., & Kuncel, N. R. (2017). Individual differences and their measurement: A review of 100 years of research. *Journal of Applied Psychology*, 254–273. https://doi.org/10.1037/apl0000151

Sasser, A. C. (2005). Gender differences in physician pay: Tradeoffs between career and family. *The Journal of Human Resources*, *40*, 477–504.

Sayer, L. C., Bianchi, S. M., & Robinson, J. P. (2004). Are parents investing less in children? Trends in mothers' and fathers' time with children. *The American Journal of Sociology*, *110*, 1–43. https://doi.org/10.1086/386270

Schulze, U. (2015). The gender wage gap among PhDs in the UK. *Cambridge Journal of Economics*, *39*, 599–629. https://doi.org/10.1093/cje/bev001

Simonton, D. K. (2014). *The Wiley handbook of genius*. Wiley Blackwell.

Snyder, T. D., de Brey, C., & Dillow, S. A. (2019). *Digest of education statistics, 2017*.

Stanley, J. C. (1977). Rationale of the Study of Mathematically Precocious Youth (SMPY) during its first five years of promoting educational acceleration. In J. C. Stanley, W. C. George, & C. H. Solano (Eds.), *The gifted and the creative: A fifty-year perspective* (pp. 75–112). Johns Hopkins University Press.

Stanley, J. C. (1990). Leta Hollingworth's contributions to above-level testing of the gifted. *Roeper Review*, *12*, 161–171.

Sterling, A. D., Thompson, M. E., Wang, S., Kusimo, A., Gilmartin, S., & Sheppard, S. (2020). The confidence gap predicts the gender pay gap among STEM graduates. *Proceedings of the National Academy of Sciences*, *117*, 30303–30308. https://doi.org/10.1073/pnas.2010269117

Stewart-Williams, S., & Halsey, L. G. (2021). Men, women and STEM: Why the differences and what should be done? *European Journal of Personality*, *35*, 3–39. https://doi.org/10.1177/0890207020962326

Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in science, technology,

engineering, and mathematics education. *Psychological Science*, *29*, 581–593. https://doi.org/10.1177/0956797617741719

Stoet, G., & Geary, D. C. (2020). Gender differences in the pathways to higher education. *Proceedings of the National Academy of Sciences*, *117*, 14073–14076. https://doi.org/10.1073/pnas.2002861117

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in Interests. *Psychological Bulletin*, *135*, 859–884. https://doi.org/10.1037/a0017364

Taniguchi, H. (1999). The timing of childbearing and women's wages. *Journal of Marriage and Family*, *61*, 1008–1019. https://doi.org/10.2307/354020

Taylor, H. C., & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. *Journal of Applied Psychology*, *23*, 565–578. https://doi.org/10.1037/h0057079

Terman, L. M. (1926). *Genetic studies of genius: Vol. I. Mental and physical traits of a thousand gifted children* (2nd ed.). Stanford University Press.

Terman, L. M., & Oden, M. H. (1959). *Genetic studies of genius. Vol. V. The gifted group at mid-life.* Stanford University Press.

Thorndike, R. L. (1949). *Personnel selection: Test and measurement techniques*. John Wiley & Sons.

Tinsley, H. E. (1993). Special issue on the Theory of Work Adjustment. *Journal of Vocational Behavior*, *43*, 1–4. https://doi.org/10.1006/jvbe.1993.1025

Todd, E. L. (2001). *Educational attainment and family gaps in women's wages: Evidence from five industrialized countries* (No. 246; Luxembourg Income Study).

Tyler, L. E. (1992). Counseling psychology: Why? *Professional Psychology: Research and Practice*, *23*, 342–344. https://doi.org/10.1037/0735-7028.23.5.342

U.S. Department of Labor. (2009). *An analysis of the reasons for the disparity in wages between men and women*. https://www.shrm.org/hr-today/public-policy/hr-public-policy-issues/Documents/Gender Wage Gap Final Report.pdf

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence*, *38*, 412–423. https://doi.org/10.1016/j.intell.2010.04.006

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

Waldfogel, J. (1997). The effect of children on women's wages. *American Sociological Review*, *62*, 209–217. https://doi.org/10.2307/2657300

Warne, R. T. (2012). History and development of above-level testing of the gifted. *Roeper Review*, *34*, 183–193. https://doi.org/10.1080/02783193.2012.686425

Weeden, K. A., Cha, Y., & Bucca, M. (2016). Long work hours, part-time work, and trends in the gender gap in pay, the motherhood wage penalty, and the fatherhood wage premium. *RSF: Russell Sage Foundation Journal of the Social Sciences*, *2*, 71–102. https://doi.org/10.7758/rsf.2016.2.4.03

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*, 1686. https://doi.org/10.21105/joss.01686

Wilde, E. T., Batchelder, L., & Ellwood, D. T. (2010). *The mommy track divides: The impact of childbearing on wages of women of differing skill levels* (No. 16582; NBER Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w16582

Williams, J. C., & Boushey, H. (2010). *The three faces of work-family conflict: The poor, the professionals, and the missing middle*. https://community-wealth.org/sites/clone.community-wealth.org/files/downloads/paper-williams-boushey.pdf

Williams, W. M., & Ceci, S. J. (2012). When scientists choose motherhood. *American Scientist*, *100*, 138–145. https://doi.org/10.1511/2012.95.138

Williams, W. M., & Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Science*, *112*, 5360–5365. https://doi.org/10.1073/pnas.1418878112

Williamson, E. G. (1965). *Vocational counseling: Some historical, philosophical, and theoretical perspectives*. McGraw-Hill.

Wolfinger, N. H., Mason, M. A., & Goulden, M. (2008). Problems in the pipeline: Gender, marriage, and fertility in the ivory tower. *The Journal of Higher Education*, *79*, 388–405. https://doi.org/10.1353/jhe.0.0015

Yu, W., & Kuo, J. C.-L. (2017). The motherhood wage penalty by work conditions: How do occupational characteristics hinder or empower mothers? *American Sociological Review*, *82*, 744–769. https://doi.org/10.1177/0003122417712729

Zakaria, F. (2008). *The post-American world*. W. W. Norton & Company.

Zuckerman, H. (1977). *Scientific elite*. Free Press.