Evaluating the Methodology of Social Science Research on Sexual Orientation and Parenting: A Tale of Three Studies

Gregory M. Herek[*]

This Article evaluates the validity and generalizability of findings from three studies that have been cited as evidence that children are negatively affected by having parents who are members of a same-sex couple, or are lesbian, gay, or bisexual. I begin by summarizing key findings from empirical research on sexual minority parenting and families. This is followed by a discussion of sexual stigma, which defines the cultural context in which the parenting studies have been conceived, conducted, interpreted, and applied to legal and policy questions. Next I explain three general methodological considerations in evaluating social science research, including studies of sexual orientation and parenting: how variables are defined and measured, how samples are created, and how researchers take into account the effects of extraneous factors that may be sources of group

 Professor of Psychology, University of California at Davis. B.A., University of Nebraska at Omaha, 1977; Ph.D., University of California at Davis, 1983; Post-doctoral Fellow, Yale University, 1983–85. This paper is based on a presentation at the UC Davis Law Review's *Not Equal Yet* Symposium on February 7, 2014. I acknowledge a great intellectual debt to the psychologists and attorneys with whom I have had the privilege of interacting and collaborating over the past three decades in preparing amicus briefs submitted by the American Psychological Association and other professional associations in cases addressing the rights of sexual minorities. Those cases include *Bowers v. Hardwick* (1986), *Romer v. Evans* (1996), *Lawrence v. Texas* (2003), *In re Marriage Cases* (2008), and *United States v. Windsor* (2013). Many of the ideas discussed in this paper were proposed and developed in those briefs. At the risk of omitting individuals who played important roles in this process, I express special thanks to Clinton Anderson, Anne Peplau, Charlotte Patterson, Nathalie Gilfoyle, Paul Smith, Bill Hohengarten, and David Ogden. I learned from and was inspired by all of them. I also thank Deborah Goldfarb for her helpful comments on an earlier version of this paper. Address correspondence to the author, Department of Psychology, University of California, Davis, CA 95616-8686.

differences. Applying these considerations, I demonstrate that studies by Sarantakos, Regnerus, and Allen all suffer from one or more serious methodological flaws whose effect is to make each study marginal or completely irrelevant to empirically-based discussions of parenting and sexual orientation. The paper concludes by returning briefly to discussing the role that stigma plays in framing societal reactions to differences — whether actual or supposed — in the adjustment and well-being of children raised in different family structures, and the benefit that legal recognition of family relationships confers on children, regardless of their parents' sexual orientation.

TABLE OF CONTENTS

INTRODUCTION

For at least four decades, questions about the well-being of children raised by sexual minority individuals or same-sex couples have figured prominently in public debates about adoption, foster parenting, child custody and visitation, and marriage equality.[1] Those debates have often focused on questions of law, personal values, and morality, but many have also included empirically testable assertions about the relationship (or lack thereof) between parents' sexual orientation and their children's developmental outcomes. The social science research that has empirically assessed the validity of those assertions is the focus of the present paper.[2]

More specifically, I have two principal aims. First, I describe some of the methodological criteria that social scientists apply when evaluating the validity of empirical research in this area. In doing so, I also explain key concepts and terminology for the benefit of readers who are not themselves trained in research methods. Second, I assess the validity and generalizability of findings from three studies that have been cited to support the argument that children are negatively affected by having parents who are lesbian, gay, or bisexual individuals, or members of a same-sex couple.[3] I demonstrate that each study suffers from one or more serious methodological flaws whose effect is to make it marginal

[1] Charlotte J. Patterson, *Children of Lesbian and Gay Parents: Psychology, Law and Policy*, 64 AM. PSYCHOLOGIST 727, 727 (2009) [hereinafter *Children of Lesbian and Gay Parents*]. *See generally* COURTNEY G. JOSLIN, SHANNON MINTER & CATHERINE SAKIMURA, LESBIAN, GAY, BISEXUAL AND TRANSGENDER FAMILY LAW § 1:1 (2014) (noting that many same-sex couples have a hard time prevailing in child custody cases); Donna Hitchens, *Social Attitudes, Legal Standards and Personal Trauma in Child Custody Cases*, 5 J. HOMOSEXUALITY 89, 93-94 (1979) (noting that judges probe deeper into a parent's sexual life and relationships in child custody cases if the parent is homosexual rather than heterosexual).

[2] Throughout this article, I use the term "research" to refer to empirical studies of behavioral and social phenomena that utilize the scientific method, i.e., systematic observation and measurement that can be replicated by other researchers with appropriate training. I frequently use the descriptor "social science" for that research to highlight the fact that it includes contributions from multiple disciplines, including psychology, sociology, anthropology, and political science.

[3] *See* Douglas W. Allen, *High School Graduation Rates Among Children of Same-Sex Households*, 11 REV. ECON. HOUSEHOLD 635, 635 (2013) [hereinafter *High School Graduation Rates*]; Mark Regnerus, *How Different Are the Adult Children of Parents Who Have Same-Sex Relationships? Findings from the New Family Structures Study*, 41 SOC. SCI. RES. 752, 764, 766 (2012) [hereinafter *How Different*]; Sotirios Sarantakos, *Children in Three Contexts: Family, Education, and Social Development*, 21 CHILD. AUSTL. 23, 23 (1996).

or completely irrelevant to empirically-based discussions of parenting and sexual orientation.

The paper begins with a brief summary of key findings from empirical research on sexual minority parenting and families.[4] This is followed by a discussion of sexual stigma, a defining feature of the cultural context in which these studies have been conceived, conducted, interpreted, and applied to legal and policy questions.[5] After providing this background, I explain three general methodological considerations relevant to evaluating research on sexual orientation and parenting: how variables are defined and measured, how samples are created, and how researchers take into account the effects of extraneous factors that might cause groups to differ on key variables.[6] Applying these considerations, I identify critical methodological flaws in studies authored by Sarantakos, Regnerus, and Allen.[7] The paper concludes by returning briefly to a discussion of the role that stigma plays in framing societal reactions to differences — whether actual or supposed — in the adjustment and well-being of children raised in different family structures, and the benefit that legal recognition of family relationships confers on all children, regardless of their parents' sexual orientation.

I. BACKGROUND

A. Children of Sexual Minority Parents: A Brief Summary of the Research

The present paper is not intended to provide a thorough review of empirical research assessing whether and how children's well-being is affected by their parents' sexual orientation.[8] As a foundation for the discussion that follows, however, a general familiarity with the main findings from research in this area is necessary. This background is important because social science, like all science, is a cumulative

[4] *See infra* Part I.A.

[5] *See infra* Part I.B.

[6] *See infra* Part I.C.

[7] *See infra* Part II.

[8] For literature reviews and more extensive discussions of this topic, see, for example, ABBIE E. GOLDBERG, LESBIAN AND GAY PARENTS AND THEIR CHILDREN: RESEARCH ON THE FAMILY LIFE CYCLE (2010); Wendy D. Manning, Marshal Neal Fettro & Esther Lamidi, *Child Well-Being in Same-Sex Parent Families: Review of Research Prepared for American Sociological Association Amicus Brief*, 33 POPULATION RES. & POL'Y REV. 485, 485 (2014); Patterson, *Children of Lesbian and Gay Parents*, *supra* note 1; Charlotte J. Patterson, *Family Lives of Lesbian and Gay Adults*, *in* THE HANDBOOK OF MARRIAGE AND THE FAMILY 659, 659 (G.W. Peterson & K.R. Bush eds., 2013) [hereinafter *Family Lives*].

enterprise. The systematic replication of researchers' observations and measurements — by the original researchers themselves and by others with appropriate training — is at the core of the scientific method.[9] The findings of any single study might result from sampling anomalies, unknown extraneous variables, or other factors. But when a reliable (i.e., consistent) pattern of findings is observed across multiple studies that have been competently conducted by different researchers with different samples, greater confidence can be placed in those findings. Thus, such patterns, to the extent that they have been observed, provide a context for evaluating each new study.

Before focusing on questions of sexual orientation and parenting, it is appropriate to consider what is known about child development in general. Three factors have been shown to be especially important in fostering children's healthy adjustment.[10] Children generally benefit when: (a) they have warm and stable relationships with their parents or parent figures; (b) their home environment affords them adequate economic, social, and physical resources, thus protecting them from poverty and social isolation; and (c) the significant adults in their lives have harmonious, supportive relationships with each other.[11] Conflicts

[9] *See generally, e.g.*, KARL R. POPPER, THE LOGIC OF SCIENTIFIC DISCOVERY 23 (Routledge 2002) (1935) ("Only when certain events recur in accordance with rules or regularities, as is the case with repeatable experiments, can our observations be tested — in principle — by anyone. We do not take even our own observations quite seriously, or accept them as scientific observations, until we have repeated and tested them. Only by such repetitions can we convince ourselves that we are not dealing with a mere isolated 'coincidence', but with events which, on account of their regularity and reproducibility, are in principle inter-subjectively testable.").

[10] *Adjustment* is "a broad umbrella term referring to personal characteristics (including the absence of psychological or psychiatric symptoms and the absence of behavior problems) that allow children, adolescents, and adults to function well in their everyday life. Well-adjusted individuals have sufficient social skills to get along with others (at school, in social settings, and at work), to get along and comply with rules and authority, to function well at school and in the workplace, and to establish and maintain meaningful intimate relationships." Michael E. Lamb, *Mothers, Fathers, Families, and Circumstances: Factors Affecting Children's Adjustment*, 16 APPLIED DEVELOPMENTAL SCI. 98, 99 (2012).

[11] For literature reviews that discuss these three factors, see *id.* and Charlotte J. Patterson & Paul D. Hastings, *Socialization in the Context of Family Diversity*, *in* HANDBOOK OF SOCIALIZATION: THEORY AND RESEARCH 328, 330-31 (Joan E. Grusec & Paul D. Hastings eds., 2007). *See also* Daniel Potter, *Same-Sex Parent Families and Children's Academic Achievement*, 74 J. MARRIAGE & FAM. 556, 567 (2012) (regarding the importance of household stability); Michael J. Rosenfeld, *Nontraditional Families and Childhood Progress Through School*, 47 DEMOGRAPHY 755, 755 (2010) [hereinafter *Nontraditional Families*] (regarding the importance of socioeconomic factors); Jennifer L. Wainright & Charlotte J. Patterson, *Delinquency, Victimization, and Substance Use Among Adolescents with Female Same-Sex Parents*, 20 J. FAM. PSYCHOL. 526, 528 (2006)

between parental figures that lead to family instability, household disruption, and divorce are often associated with poorer adjustment and problems that can last into adulthood.[12]

Starting from these general observations, we can ask how parents' sexual orientation is related to their children's well-being. The research addressing this question has yielded notably consistent findings despite variability across the published studies in sampling strategies, measurement techniques, and other aspects of methodology. All else being equal, comparisons across multiple indicators of mental health and social adjustment reveal few, if any, reliable disparities between heterosexual and sexual minority families. These indicators include: children's levels of self-esteem, anxiety, and depression; their levels of hyperactivity, unsociability, emotional difficulty, and conduct difficulty; and their performance in social arenas such as school, sports, and friendships.[13]

Most research has focused on children raised by female couples or lesbian single mothers. However, the literature increasingly includes studies that compare children raised by gay fathers to children from households headed by heterosexual or lesbian parents. Those studies similarly fail to find significant differences in children's well-being across family types.[14] As a noted developmental psychologist summarized the empirical research findings:

[hereinafter *Delinquency*]; Jennifer L. Wainright & Charlotte J. Patterson, *Peer Relations Among Adolescents with Female Same-Sex Parents*, 44 DEVELOPMENTAL PSYCHOL. 117, 123 (2008) [hereinafter *Peer Relations*]; Jennifer L. Wainright, Stephen T. Russell & Charlotte J. Patterson, *Psychosocial Adjustment, School Outcomes, and Romantic Relationships of Adolescents with Same-Sex Parents*, 75 CHILD DEV. 1886, 1893-95 (2004) (regarding the importance of the quality of relationships between parent and child, and children's perception that adults, teachers, and friends care about them).

[12] For a discussion on the negative effects of household disruption and parental divorce, see Paul R. Amato, *Children of Divorce in the 1990s: An update of the Amato and Keith (1991) Meta-analysis*, 15 J. FAM. PSYCHOL. 355, 355 (2001) and Paul R. Amato & Bruce Keith, *Parental Divorce and the Well-being of Children: A Meta-analysis*, 110 PSYCHOL. BULL. 26, 26 (1991).

[13] For reviews of this research literature, see GOLDBERG, *supra* note 8, at 125-44; Patterson, *Children of Lesbian and Gay Parents*, *supra* note 1, at 731-32; Patterson, *Family Lives*, *supra* note 8, at 668-71; Judith Stacey & Timothy J. Biblarz, *(How) Does the Sexual Orientation of Parents Matter?*, 66 AM. SOC. REV. 159, 169 (2001).

[14] *See* Stephen Erich, Patrick Leung & Peter Kindle, *A Comparative Analysis of Adoptive Family Functioning with Gay, Lesbian, and Heterosexual Parents and their Children*, 1 J. GLBT FAM. STUD. 43, 45 (2005); Rachel H. Farr, Stephen L. Forssell & Charlotte J. Patterson, *Parenting and Child Development in Adoptive Families: Does Parental Sexual Orientation Matter?*, 14 APPLIED DEVELOPMENTAL SCI. 164, 174 (2010); Justin A. Lavner, Jill Waterman & Letitia Anne Peplau, *Can Gay and Lesbian Parents Promote Healthy Development in High-Risk Children Adopted From Foster Care?*, 82 AM.

> [T]he adjustment of children and adolescents is best accounted for by variations in the quality of the relationships with their parents, the quality of the relationship between the parents or significant adults in the children's and adolescent's lives, and the availability of economic and socio-economic resources. . . . The parents' sex and sexual orientation, like other characteristics of family structure, do not affect either the capacity to be good parents or their children's healthy development.[15]

Another leading researcher in this area commented on the "remarkably consistent" results obtained in studies of children and adolescents alike, regardless of which outcome variables were measured: "In study after study, the offspring of lesbian and gay parents have been found to be at least as well adjusted overall as those of other parents."[16]

To say that parents' sexual orientation has not been found to affect children's healthy development is not to assert that no differences have been observed between children raised by heterosexuals and those raised by lesbians, gay men, or bisexuals. However, few reports of statistically significant differences in the outcomes of children from different family structures (e.g., headed by lesbian parents vs. a heterosexual couple) have been replicated by other researchers. In the absence of such replications, we cannot dismiss the possibility that a difference observed in only one or two studies — regardless of whether it appears to favor the children of same-sex or different-sex couples — may be an artifact of the study sample or methodology, and not characteristic of the population at large.

B. The Context: Sexual Stigma

Controversies and debates about the social science research on sexual orientation and parenting must be understood within their cultural context, which is defined by "stigma." Whereas numerous definitions have been offered for the term stigma, they tend to converge in

J. ORTHOPSYCHIATRY 465, 465 (2012); *see also* Susan Golombok et al., *Adoptive Gay Father Families: Parent-Child Relationships and Children's Psychological Adjustment*, 85 CHILD DEV. 456, 456 (2014) [hereinafter *Adoptive Gay Father Families*]. For an earlier review of the then-extant research on gay fathers and their children, see generally Charlotte J. Patterson, *Gay Fathers*, *in* THE ROLE OF THE FATHER IN CHILD DEVELOPMENT 397, 402 (Michael E. Lamb ed., 2004).

15 Lamb, *supra* note 10, at 106.

16 Patterson, *Children of Lesbian and Gay Parents*, *supra* note 1, at 732.

highlighting several key ideas. Stigma is generally used to refer to culturally shared knowledge about society's negative regard for and according of inferior status to members of a particular group or category. Research on stigma has usually focused on conditions, statuses, and attributes that are enduring rather than transient, and are a source of disadvantage and disempowerment.[17] The sociologist Erving Goffman, perhaps the most famous social scientist to write about stigma, characterized it as "an undesired differentness" whose significance and status derive less from the devalued trait or group membership itself than from the socially constructed meanings attached to it by the larger culture.[18]

"Sexual stigma" refers to the negative regard, inferior status, and relative powerlessness that society collectively accords any nonheterosexual behavior, identity, relationship, or community.[19] Like other forms of stigma, it is fundamentally about power.[20] Stigma-based differences in power and status are legitimated and perpetuated by society's institutions and ideological systems in the form of "structural stigma," which "is formed by sociopolitical forces and represents the policies of private and governmental institutions that restrict the opportunities of stigmatized groups."[21] The legitimacy of structural sexual stigma — in the form of laws and institutional policies that treat sexual minority individuals and same-sex couples differently from heterosexuals — is being challenged with increasing success in many countries. Examples from the United States and elsewhere include changes in military policies to permit service by openly lesbian, gay, and bisexual personnel, as well as the legal recognition of same-sex relationships. At the same time, structural sexual stigma has recently been reinforced in other countries by legislation that attaches severe

[17] For a more detailed discussion of this definition of stigma, see Gregory M. Herek, *Sexual Stigma and Sexual Prejudice in the United States: A Conceptual Framework*, 54 NEB. SYMP. ON MOTIVATION 65, 66 (2009) [hereinafter *Sexual Stigma*]. For other discussions of stigma, see generally ERVING GOFFMAN, STIGMA: NOTES ON THE MANAGEMENT OF SPOILED IDENTITY 2-4 (1963); EDWARD E. JONES ET AL., SOCIAL STIGMA: THE PSYCHOLOGY OF MARKED RELATIONSHIPS 9 (1984); and Jennifer Crocker et al., *Social Stigma*, *in* THE HANDBOOK OF SOCIAL PSYCHOLOGY 504-05 (Daniel T. Gilbert et al. eds., 4th ed. 1998).

[18] GOFFMAN, *supra* note 17, at 5; *see also* Crocker et al., *supra* note 17, at 504; Herek, *Sexual Stigma*, *supra* note 17, at 66-67.

[19] Herek, *Sexual Stigma*, *supra* note 17, at 67.

[20] *See id.* at 66-67; *see also* Bruce G. Link & Jo C. Phelan, *Conceptualizing Stigma*, 27 ANN. REV. SOC. 363, 375-76 (2001).

[21] Patrick W. Corrigan et al., *Structural Stigma in State Legislation*, 56 PSYCHIATRIC SERVS. 557, 557 (2005). For further discussion of structural stigma, see generally Link & Phelan, *supra* note 20, at 372-73 (discussing structural discrimination and institutionalized stigmas).

penalties — sometimes death — to sexual behavior between two adults of the same sex and effectively prohibits public advocacy on behalf of sexual and gender minorities.[22]

Structural sexual stigma has also been referred to as "heterosexism."[23] As a core component of society's institutions, heterosexism ensures that sexual minorities as a group are less powerful than heterosexuals.[24] One way it accomplishes this is by fostering a "heterosexual assumption" throughout society whereby all people are presumed to be heterosexual, and heterosexual behavior and different-sex relationships are considered natural and unproblematic. When gay, lesbian, and bisexual people overcome the invisibility imposed by this assumption, their sexuality is viewed as abnormal, unnatural, and requiring explanation. Differences observed between them and heterosexuals are generally interpreted as evidence of deficits on their part.[25] By legitimating and reinforcing sexual minorities' undesired differentness, relative powerlessness, and inferior status relative to heterosexuals, heterosexism creates a context that fosters individual enactments of

---

[22] *See, e.g.*, David M. Herszenhorn, *Gays in Russia Find No Haven, Despite Support from the West*, N.Y. TIMES (Aug. 11, 2013), *available at* http://www.nytimes.com/2013/08/12/world/europe/gays-in-russia-find-no-haven-despite-support-from-the-west.html (discussing Russian legislation that aims to suppress homosexuality and Russia's fledgling gay rights movement); *Homophobia and Punitive Laws Continue to Threaten HIV Responses and Human Rights*, UNAIDS (Aug. 28, 2012), http://www.unaids.org/en/resources/presscentre/featurestories/2012/august/20120828punitivelaws/ (discussing the negative impact that antigay discrimination and laws criminalizing same-sex sexual behavior have on HIV education, prevention, and treatment); *Mapping Anti-Gay Laws in Africa*, AMNESTY INT'L (Jan. 13, 2014), http://www.amnesty.org.uk/lgbti-lgbt-gay-human-rights-law-africa-uganda-kenya-nigeria-cameroon (listing African countries with laws criminalizing same-sex sexual behavior or relationships); David Smith, *Chad Becomes 37th African State to Seek Ban on Homosexuality*, THE GUARDIAN (Sept. 22, 2014), http://www.theguardian.com/world/2014/sep/22/chad-37th-african-state-seeking-ban-homosexuality (describing enactment of a law criminalizing same-sex sexual behavior). Concerning international differences in individuals' views of whether homosexuality should be accepted or rejected by society, see generally *The Global Divide on Homosexuality*, PEW RESEARCH CTR. (June 4, 2013), http://www.pewglobal.org/files/2013/06/Pew-Global-Attitudes-Homosexuality-Report-FINAL-JUNE-4-2013.pdf.

[23] "Heterosexism" has been defined in a variety of ways; in the present paper, it refers to structural sexual stigma's underlying ideology and its various manifestations. *See* Herek, *Sexual Stigma*, *supra* note 17, at 67.

[24] *Id.*

[25] *See* Gregory M. Herek, *Sexual Orientation Differences as Deficits: Science and Stigma in the History of American Psychology*, 5 PERSP. ON PSYCHOL. SCI. 693, 693-94 (2010); *see also, e.g.*, Diana Baumrind, *Commentary on Sexual Orientation: Research and Social Policy Implications*, 31 DEVELOPMENTAL PSYCHOLOGY 130, 133-35 (1995); Stacey & Biblarz, *supra* note 13, at 162.

antigay stigma, including ostracism, harassment, discrimination, and violence.[26]

One manifestation of sexual stigma is institutionalized discrimination against sexual minority parents, e.g., the historic lack of legal recognition for their family relationships and the obstacles imposed on them in divorce and custody proceedings.[27] Such discrimination has often been justified by claims that children are harmed by having lesbian, gay, or bisexual parents whereas married heterosexual parents are the "gold standard" for healthy and normal families.[28] Within this context, any observed differences between heterosexuals and sexual minorities are interpreted as evidence of deficits in the latter group.

Empirical research on sexual orientation and parenting has been conducted and publicly discussed against this backdrop of stigma and the assumption that differences indicate deficits. For example, the mental health profession considers homosexuality and bisexuality to be natural variants of human sexuality.[29] Nevertheless, some individuals and groups regard homosexuality in strongly negative terms. They promote the belief that parents should try to prevent homosexuality in their children and that gay and lesbian people can and should change

---

[26] For a discussion of sexual stigma's manifestations by individuals, e.g., in the form of heterosexuals' prejudice against sexual minorities, see generally Herek, *Sexual Stigma*, *supra* note 17, at 74-98.

[27] *See* JOSLIN, MINTER & SUKIMURA, *supra* note 1, § 8:1.

[28] This argument is often buttressed by the claim that scientific research has established that a family headed by married biological parents is the optimal environment for child rearing. *See, e.g.*, *Ten Arguments from Social Science Against Same-Sex Marriage*, FAMILY RESEARCH COUNCIL (last visited Nov. 3, 2014), http://www.frc.org/get.cfm?i=if04g01 (citing McLanahan & Sandefur to make this argument). The body of research cited in these arguments, however, does not include studies comparing children according to their parents' sexual orientation. Rather, it comprises studies comparing children's outcomes according to the number of parents who raised them, i.e., one or two. It has tended to show that children have fewer problems when they are raised by married heterosexual parents than by a mother who never married or a parent who was single due to divorce, separation, or the death of a spouse. For a general review of that literature, see SARA MCLANAHAN & GARY SANDEFUR, GROWING UP WITH A SINGLE PARENT: WHAT HURTS, WHAT HELPS 19 (1994). *See also* Leonard M. Lopoo & Thomas DeLeire, *Family Structure and the Economic Wellbeing of Children in Youth and Adulthood*, 43 SOC. SCI. RES. 30, 41 (2014) (showing that many of the negative outcomes of children raised by a single parent may be explained by economic disadvantages). However, equivalent studies comparing children reared by same-sex couples with those raised by a single lesbian, gay, or bisexual parent have not yet appeared in the research literature.

[29] *See* Brief of The American Psychological Association et al. as Amici Curiae Supporting Respondents at 7-10, United States v. Windsor, 133 S. Ct. 2675 (2013) (No. 12-307).

their orientation to heterosexual.[30] From such a perspective comes an assumption that if children raised by non-heterosexual parents were found to be more likely than other children to grow up to be lesbian, gay, or bisexual, this would be a significant problem. Similarly, according to the differences-are-deficits ideology, if the children of non-heterosexual parents were observed to be less conforming to traditional gender roles than the children of heterosexuals, this would indicate that sexual minority parents are less capable.

In fact, few studies have been published to date that examine the sexual orientation of adults who were raised by parents in a same-sex committed relationship, probably because children in the cohort of what has been called the lesbian and gay "baby boom" of the 1990s are only now reaching adulthood.[31] However, the available empirical data are consistent with the conclusion that the vast majority of those children eventually grow up to be heterosexual.[32] As for gender attitudes and behavior, most published studies comparing the children of lesbian and heterosexual mothers have found no reliable differences in adherence to cultural norms defining feminine and masculine behavior.[33] However, some have found that children raised by lesbian

[30] *See generally* BOB DAVIES & LORI RENTZEL, COMING OUT OF HOMOSEXUALITY: NEW FREEDOM FOR MEN AND WOMEN (1993) (prescribing religiously based strategies for changing one's homosexual orientation to heterosexual); JOSEPH NICOLOSI & LINDA AMES NICOLOSI, A PARENT'S GUIDE TO PREVENTING HOMOSEXUALITY (2002) (purporting to provide guidance for parents who want to prevent their child from being homosexual). Regarding efforts to change sexual orientation, see generally ARIEL SHIDLO ET AL., SEXUAL CONVERSION THERAPY: ETHICAL, CLINICAL, AND RESEARCH PERSPECTIVES (2001).

[31] Although population estimates are not available, the number of children being raised from birth by female-female and male-male couples appeared to increase substantially in the 1990s. This phenomenon is often referred to as a "baby boom." Charlotte J. Patterson, *Families of the Lesbian Baby Boom: Parents' Division of Labor and Children's Adjustment*, 31 DEVELOPMENTAL PSYCHOL. 115, 115 (1995).

[32] One study using data from the National Longitudinal Lesbian Family Study to examine sexuality in seventeen-year olds reared by lesbian parents found that two of the thirty-seven boys (but none of the thirty-seven girls) reported they were predominantly or exclusively homosexual. Most of the teens (thirty girls and thirty-four boys) reported they were predominantly or exclusively heterosexual. Nanette K. Gartrell et al., *Adolescents of the U.S. National Longitudinal Lesbian Family Study: Sexual Orientation, Sexual Behavior, and Sexual Risk Exposure*, 40 ARCHIVES OF SEXUAL BEHAV. 1199, 1204 (2011); *see also* GOLDBERG, *supra* note 8, at 132-34; Patterson, *Family Lives*, *supra* note 8, at 669-70.

[33] *See, e.g.*, Henny Bos & Theo G. M. Sandfort, *Children's Gender Identity in Lesbian and Heterosexual Two-Parent Families*, 62 SEX ROLES 114, 120, 122 (2010) (Finding that Dutch children raised by lesbian or heterosexual couples did not differ significantly on gender typicality, contentedness with their gender, or felt pressure from peers to conform to gender stereotypes, although children in lesbian families were less likely to regard their own gender as superior or to feel parental pressure to conform to gender

mothers or same-sex couples are more accepting of gender nonconformity in others and are more flexible in their own patterns of gender-role behaviors (e.g., during play).[34] Within the ideological framework of heterosexism, this difference is likely to be seen as a deficit, even though holding flexible attitudes about gender roles — for example, girls aspiring to traditionally masculine occupations such as astronaut or engineer — can instead be considered psychologically healthy for children.[35]

In summary, heterosexism creates a set of implicit ground rules for interpreting scientific research on parenting by same-sex couples and lesbian, gay, and bisexual individuals. A presumption exists that non-heterosexual family forms are bad for children unless scientific research consistently fails to find differences related to parent sexual orientation. Moreover, application of the differences-are-deficits model is not limited to outcomes such as psychological and social adjustment. Departures from the gold standard of families headed by married heterosexuals are equated with deficits even when such differences would be benign (e.g., if children raised by same-sex couples showed an increased likelihood of growing up to be non-heterosexual) or potentially adaptive (e.g., if they grew up to be generally more tolerant of diverse social groups or flexible in their attitudes concerning gender roles).

As noted in the previous section, research on children raised by sexual minority parents has not revealed reliable differences in adjustment according to parent sexual orientation. Thus, this body of scientific literature poses a challenge to the legitimacy of sexual stigma related to parenting and family structure. It is not difficult to understand

stereotypes; Farr, Forssell & Patterson, *supra* note 14, at 175 (finding that young children adopted in infancy by lesbian and gay parents showed typical gender development). For literature reviews, see GOLDBERG, *supra* note 8, at 130-32 and Patterson, *Family Lives*, *supra* note 8, at 668-70.

[34] *See* Megan Fulcher et al., *Individual Differences in Gender Development: Associations with Parental Sexual Orientation, Attitudes, and Division of Labor*, 58 SEX ROLES 330, 336 (2008); Abbie E. Goldberg et al., *Gender-Typed Play Behavior in Early Childhood: Adopted Children with Lesbian, Gay, and Heterosexual Parents*, 67 SEX ROLES 503, 511 (2012); Richard Green et al., *Lesbian Mothers and Their Children: A Comparison with Solo Parent Heterosexual Mothers and Their Children*, 15 ARCHIVES OF SEXUAL BEHAV. 167, 179 (1986); Erin L. Sutfin et al., *How Lesbian and Heterosexual Parents Convey Attitudes About Gender to Their Children: The Role of Gendered Environments*, 58 SEX ROLES 501, 508 (2008).

[35] *See, e.g.*, Anne E. Barrett & Helene Raskin White, *Trajectories of Gender Role Orientations in Adolescence and Early Adulthood: A Prospective Study of the Mental Health Effects of Masculinity and Femininity*, 43 J. HEALTH & SOC. BEHAV. 451, 462 (2002) (finding that higher levels of masculinity throughout adolescence are associated with fewer depressive symptoms in early adulthood for both girls and boys).

why new empirical studies that appear to call into question the validity of this body of research would receive a considerable amount of attention.

Two such studies have been recently published and they — along with a third, older study that also reports anomalous findings — have been introduced in legal proceedings concerning sexual minority couples and parenting.[36] As with all empirical research, their credibility depends on their methodological quality. No study is methodologically perfect, and the impact of methodological limitations on the substantive conclusions that can be drawn from a study varies. Some problems merely necessitate that a study's conclusions be appropriately qualified whereas others completely invalidate the study. Identifying such limitations and assessing their importance require a working knowledge of the criteria that social scientists use to assess the methodological quality of empirical research. I explain some of these criteria in the next section. Then I demonstrate that each of the three studies to be critically examined has one or more important methodological limitations that make it largely or entirely inapplicable to debates about parenting and sexual orientation.

C. *Evaluating Empirical Research: Methodological Considerations*

In this section, I discuss three methodological considerations that are relevant to evaluating empirical research on sexual orientation and parenting: (1) how a study's variables are conceptually and operationally defined and the extent to which the two definitions are consistent or discrepant; (2) how its sample was selected, who was included, and how closely the sample resembles the larger population to which the study's results are generalized; and (3) whether and how any relevant extraneous factors that may affect the study's outcome were controlled or otherwise taken into account.

1. Operational vs. Conceptual Definitions of Variables

A variable is a phenomenon or construct that takes on different forms or values, that is, it *varies* in some fashion.[37] The variable may assume

[36] *E.g.*, DeBoer v. Snyder, 973 F. Supp. 2d 757, 765-68 (2014) (reflecting on expert testimonies by Loren Marks, Mark Regnerus, and Douglas Allen). The studies themselves — by Allen, Regnerus, and Sarantakos — are discussed below at length.

[37] Dozens of research methodology textbooks include discussions of social science research terminology, philosophy, and procedures. Authors may provide slightly different explanations for various concepts, or may sometimes employ different terms to refer to them, but the underlying concepts are widely understood among social

differing numerical values, for example, as with chronological age or scores on a test. Or its variation may be categorical, as with political party affiliation or geographic area of residence, or ordinal, as with rankings of colleges or sports teams.

Researchers define the variables they study in two general ways. A variable's "conceptual," or theoretical, definition is the sort of formal definition found in textbooks and dictionaries. It is general, abstract, and intended to apply across situations and contexts.[38] For example, psychological depression might be conceptually defined simply as a state of feeling sad,[39] or as a mood disorder in which feelings of sadness, loss, anger, or frustration interfere with everyday life for an extended period of time.[40] Either definition could be appropriate, depending on the research question and study goals.

In contrast to a variable's conceptual definition, its "operational" definition specifies in concrete terms how it will actually be measured or manipulated in a particular research study, that is, the operations or procedures the researcher will use.[41] Because most variables studied by social scientists can be operationally defined in multiple ways, researchers face the challenge of selecting a method that is appropriate for their research question and the study population. If a variable's operational definition does not closely correspond to its conceptual definition, the researcher is effectively placed in the position of drawing conclusions about one variable based on measurements of another variable. Such conclusions are likely to be misleading or inaccurate.

The term "validity" refers to the extent to which a variable's operationalization matches its conceptual definition. Put more simply, a procedure for measuring a variable is valid to the extent that it successfully measures the phenomenon it purports to measure.[42] For example, if Researcher A is interested in the frequency with which

scientists. The citations provided here and throughout this section are merely examples of sources. Regarding the concept of a "variable," see, for example, FRED N. KERLINGER, FOUNDATIONS OF BEHAVIORAL RESEARCH: EDUCATIONAL AND PSYCHOLOGICAL INQUIRY 33-38 (1964); ELAZAR J. PEDHAZUR & LIORA PEDHAZUR SCHMELKIN, MEASUREMENT, DESIGN, AND ANALYSIS: AN INTEGRATED APPROACH 174 (1991); BART L. WEATHINGTON ET AL., RESEARCH METHODS FOR THE BEHAVIORAL AND SOCIAL SCIENCES 69 (2010).

[38] *Cf.* KERLINGER, *supra* note 37, at 33-34 (referring to the use of concepts to define a variable as a "constitutive" definition).

[39] *See Depression Definition*, MERRIAM-WEBSTER, http://www.merriam-webster.com/dictionary/depression (last visited Oct. 27, 2014).

[40] *See* Timothy Rogge, *Depression*, MEDLINEPLUS (Oct. 9, 2014), http://www.nlm.nih.gov/medlineplus/ency/article/003213.htm.

[41] *E.g.*, KERLINGER, *supra* note 37, at 34; WEATHINGTON ET AL., *supra* note 37, at 79.

[42] *E.g.*, JUM C. NUNNALLY, PSYCHOMETRIC THEORY 75 (1967).

people experience common mood states and uses the first conceptual definition of depression stated above, she might operationally define depression in terms of participants' reports of how often, on average, they feel sadness during a typical month. By contrast, if Researcher B is using the second conceptual definition in a study of the population prevalence of different forms of mental illness, he might show participants a list of symptoms that are often associated with clinical depression and ask them to report the extent to which they have experienced each symptom during the previous week. Alternatively, he might ask participants whether a doctor has ever told them that they have depression, or he might have participants clinically assessed and diagnosed by trained mental health professionals. Depending on how they are executed, each of these operationalizations could yield valid observations.

If, however, Researcher B used Researcher A's operational definition — the simple question about sadness — his data would most likely overstate the population prevalence of depression because sadness is a commonly experienced emotion that is not unique to clinical depression. Similarly, if Researcher A operationally defined sadness using Researcher B's methods, she would underestimate the frequency with which most people feel sad because she would only be measuring the type of severe sadness that is associated with clinical depression. In both examples, the variable's operational definition of depression is discrepant from its conceptual definition; the former does not provide a valid measurement of the latter. Thus, the researcher's conclusions are likely to be inaccurate.

Measures of a variable can lack validity even when they are derived from an operational definition that is consistent with the conceptual definition. Returning to the previous example, the stigma associated with mental illness makes some people reluctant to admit that they experience symptoms of depression or have been diagnosed with it. They may intentionally give researchers inaccurate self-reports. Alternatively, an assessment technique may lack validity because it requests information that respondents are incapable of providing. For example, if a researcher asks respondents to report the total number of days during the previous year on which they felt sad, most will only be able to give a rough estimate because the question makes unreasonable demands on their memory.[43]

[43] *See* ROGER TOURANGEAU ET AL., THE PSYCHOLOGY OF SURVEY RESPONSE 86 (2000).

2. Sampling and Generalizing

Observing or interviewing every member of an entire population is usually impractical. Consequently, most empirical studies are based on data obtained from a subset of the population, that is, a "sample." Researchers use two general types of sampling, referred to as probability and nonprobability sampling.[44]

In "probability sampling," all members of the "target population" (i.e., the population to which the results will be generalized) have some chance of being selected, and the probability that any individual will be included can be mathematically computed.[45] Everyone need not have the same probability of being selected but the probability must be calculable for each person.[46] In practice, probability sampling involves a series of steps that begin with defining the target population and culminate in actually collecting the data. Because some population members inevitably are eliminated at each step, evaluating a study's findings requires knowing who was excluded over the course of the sampling process.

After defining the target population, researchers must identify the members of that population who are actually available for study — the "survey population."[47] This group routinely excludes some sectors of the target population. For example, some people may not be included because they are incapable of participating (e.g., children are typically excluded from public opinion surveys, as are severely ill and hospitalized individuals). Finding or contacting other population groups may be impractical (e.g., people who are incarcerated, active-duty military personnel stationed abroad). Other groups may be eliminated because researchers lack the financial resources or capabilities to include them. For example, because it is costly to translate questions and responses into multiple languages, national surveys in the United States routinely are restricted to population members who speak English.

---

[44] For explanations of sampling terminology and methods, see, for example, GRAHAM KALTON, INTRODUCTION TO SURVEY SAMPLING (1983); LESLIE KISH, SURVEY SAMPLING (1965); Penny S. Visser et al., *Survey Research*, *in* HANDBOOK OF RESEARCH METHODS IN SOCIAL AND PERSONALITY PSYCHOLOGY (Harry T. Reis & Charles M. Judd eds., 2000).

[45] *E.g.*, KALTON, *supra* note 44, at 7; KISH, *supra* note 44, at 20; Visser et al., *supra* note 44, at 230.

[46] In a telephone survey, for example, a member of the target population who has two telephone lines is twice as likely to be selected as a person with only one line.

[47] *E.g.*, KALTON, *supra* note 44, at 6.

Once the survey population has been specified, researchers operationally define it by creating the "sampling frame." The sampling frame can be thought of as an actual or virtual list of everyone in the survey population — their names, telephone numbers, physical addresses, e-mail addresses, or other contact information.[48] Researchers then select a subset of the sampling frame, using some type of randomized procedure.[49] The members of that subset constitute the actual sample from which researchers will collect data. Randomization ensures that those who are selected for the sample do not systematically differ from the other members of the sampling frame.

Data collection proceeds after the sample has been selected. A screening procedure is typically used for ensuring that sample members are eligible for the study (e.g., that they actually belong to the survey population).[50] Ultimately, even some sample members who meet the study's eligibility criteria will not participate. Some will be unavailable or unreachable; others will explicitly refuse. The proportion of eligible sample members who actually participate in the study is quantified as a "response rate."[51] The response rate is relevant to evaluating a study's findings to the extent that responders and nonresponders differ systematically on variables related to the research question. For example, if nonresponse to an election survey is random, that is, if responders and nonresponders do not differ in any way that is relevant to voting preferences, the study results are likely to accurately describe

[48] *E.g.*, KALTON, *supra* note 44, at 7-8; KISH, *supra* note 44, at 53. The sampling frame inevitably leaves out some members of the survey population. When it consists of telephone numbers, for example, individuals without a telephone are excluded. When it consists of physical addresses, people who are homeless or living in barracks or dormitories are excluded. The term "coverage error" refers to such discrepancies between the sampling frame and the survey population. *See, e.g.*, KISH, *supra* note 44, at 53-59 (discussing sources of problems related to the sampling frame and how they can be addressed).

[49] An example of such a process would be to assign a unique identification number to every member of the sampling frame, use a computer to generate a list of random numbers (the size of the list determined by the desired sample size), and select the members of the sampling frame whose identification number matches a number on the list.

[50] If the sampling frame consists of telephone numbers, for example, screening is likely to reveal that some numbers are nonworking or belong to a commercial business rather than a private individual. Other telephone numbers might belong to a non-English-speaking person or someone who is otherwise ineligible for the study (e.g., an election poll may screen out people who are not registered voters).

[51] Numerous formulae exist for computing the response rate. *See generally* THE AM. ASS'N FOR PUB. OP. RESEARCH, STANDARD DEFINITIONS: FINAL DISPOSITIONS OF CASE CODES AND OUTCOME RATES FOR SURVEYS 44-45 (Tom W. Smith ed., 7th ed. 2011) (detailing standardized methods for calculating response rates for different types of samples).

the target population (in this case, registered voters) even when the response rate is low. However, if the response rate is low because a particular segment of the sample (e.g., political conservatives) largely decline to participate, the survey results probably will not accurately describe voter preferences in the population at large. Researchers generally strive to maximize response rates, based on the assumption that higher rates reduce the likelihood of the sample being biased.[52]

When this set of procedures is followed, findings obtained from the sample can be appropriately generalized to the target population with a known level of precision.[53] However, the study findings may not apply to the members of groups or categories that were excluded in the steps between identifying the target population and actual data collection. Thus, it is always important to determine exactly which segments of the target population are actually represented in the final sample.

In contrast to probability sampling, "nonprobability sampling" does not give everyone in the target population a known chance of inclusion.[54] Instead of enumerating all members of the population and randomly selecting a sample, the researcher recruits participants who are available, e.g., because they volunteer, happen to be in a setting where the researcher is recruiting, or belong to the social network of another study participant who recruits them at the request of the researcher.[55] Some nonprobability samples are referred to as "haphazard samples" or "convenience samples," not because they are necessarily "convenient" for the researcher to obtain (indeed, recruitment is often a time-consuming, labor-intensive process) but because they consist only of population members who were readily available.[56]

The extent to which a nonprobability sample is representative of the target population cannot be known. Nevertheless, important

---

[52] However, this assumption may not always be correct. *See, e.g.*, Allyson L. Holbrook et al., *The Causes and Consequences of Response Rates in Surveys By the News Media and Government Contractor Survey Research Firms*, *in* ADVANCES IN TELEPHONE SURVEY METHODOLOGY 500 (James M. Lepkowski et al. eds., 2008) (reporting analyses showing that lower response rates did not inevitably reduce the quality of survey data).

[53] This level of precision is expressed as a range of values within which the variable's true value (i.e., in the population) is probably located. For example, if 45% of a probability sample supports a particular candidate, statistical calculations might indicate that the proportion of the entire population supporting the candidate is probably somewhere between 42% and 48% (i.e., 45% ± 3%). In this example, the range 42–48% is called the "confidence interval," and the value ± 3 points is referred to as the "margin of error." *E.g.*, ROBERT M. GROVES ET AL., SURVEY METHODOLOGY 97-98, 382 (2004).

[54] *E.g.*, Visser et al., *supra* note 44, at 235.

[55] The last approach is often referred to as "snowball sampling." *E.g.*, *id.* at 236.

[56] *E.g.*, *id.*

contributions to knowledge can come from such samples. For very small segments of the population, probability sampling can be prohibitively expensive because such a large number of potential respondents must be screened to yield a final sample of sufficient size. Nonprobability sampling in such a situation, by contrast, is likely to be considerably less costly. Similarly, for stigmatized groups whose members are difficult to reach through traditional probability sampling methods (e.g., because they are reluctant to disclose their status to a stranger), nonprobability sampling often may be the only feasible strategy for assembling a sample.

Some research questions can be addressed simply by documenting that a phenomenon occurs or a characteristic exists within a population. For example, the hypotheses that homosexuality is inherently associated with mental illness or that being raised by a same-sex couple is inherently damaging to children could be rejected on the basis of valid documentation of one or more instances of, respectively, well-adjusted lesbians and gay men or well-adjusted individuals who were raised by a same-sex couple. It would not matter if these case studies were derived from members of a probability or nonprobability sample. Similarly, studies of phenomena that are common to all humans (e.g., as in much biomedical research) can yield valid results based on nonprobability samples.[57]

For these and other reasons, nonprobability samples have long been the mainstay of psychological, medical, and other research. This has also been the case for studies of sexual minority parenting, most of which have utilized nonprobability samples. In recent years, however, researchers have begun to make use of large probability samples for which information about the members' sexual orientation or relationship status is available. The latter construct has been operationally defined in different ways. Some studies have directly asked participants to specify how they label themselves (gay, lesbian, bisexual, heterosexual, or something else).[58] Others have inferred

[57] To the extent that relevant characteristics differ among segments of the population, however, the generalizability of findings may be limited to those sectors of the population that were included in the sample (e.g., biomedical data from an all-male sample may not be generalizable to women).

[58] For examples of surveys conducted with national probability samples that asked respondents how they label their sexual orientation, see Gregory M. Herek et al., *Demographic, Psychological, and Social Characteristics of Self-Identified Lesbian, Gay, and Bisexual Adults in a US Probability Sample*, 7 SEXUALITY RES. & SOC. POL'Y 176, 179 (2010); Gary Gates & Frank Newport, *Special Report: 3.4% of U.S. Adults Identify as LGBT*, GALLUP (Oct. 18, 2012), http://www.gallup.com/poll/158066/special-report-adults-identify-lgbt.aspx; Frank Newport, *LGBT Americans Continue to Skew Democratic*

sample members' sexual orientation based on their self-reports of cohabiting with an adult of the same sex or another proxy variable.[59]

One study, for example, used data from the Early Childhood Longitudinal Study – Kindergarten cohort to compare the academic achievement of children growing up in various family structures.[60] Parents' sexual orientation was not directly assessed but was inferred using rosters of all household members — which included information about each person's gender and relationship to the child — to identify families headed by a same-sex couple. When the effects of significant family transitions (e.g., those related to parental divorce, separation, or death) were taken into account, children in same-sex family structures showed slightly higher achievement levels than children living with their biological mother and father, but this difference was not statistically significant.

Another study used U.S. Census data to compare educational outcomes among children residing in homes with various types of family structures.[61] The Census does not directly assess respondents' sexual orientation but does allow them to indicate that they are cohabiting with a same-sex partner.[62] Among the children and parents whose addresses remained the same for at least five years (a proxy for household stability), school progress did not differ significantly between children of married heterosexual couples and same-sex cohabiting couples when differences in parents' socioeconomic status were statistically controlled.[63] As the researcher concluded, his analyses showed that "children raised by same-sex couples have no fundamental deficits in making normal progress through school."[64]

and Liberal, GALLUP (July 30, 2014), http://www.gallup.com/poll/174230/lgbt-americans-continue-skew-democratic-liberal.aspx.

59 *E.g.*, Rosenfeld, *Nontraditional Families*, *supra* note 11, 760-61; Wainright, Russell & Patterson, *supra* note 11, at 1889-90.

60 Potter, *supra* note 11, at 556.

61 Rosenfeld, *Nontraditional Families*, *supra* note 11, at 761-63; *see also* Michael J. Rosenfeld, *Reply to Allen et al.*, 50 DEMOGRAPHY 963, 963-64 (2012) [hereinafter *Reply to Allen*].

62 Rosenfeld, *Nontradtional Families*, *supra* note 11, at 760-61.

63 *Id.* at 766-69.

64 Rosenfeld, *Nontraditional Families*, *supra* note 11, at 772; Rosenfeld, *Reply to Allen*, *supra* note 61, at 965-66 (showing that differences between children raised by same-sex couples and children from other family types become nonsignificant when sociodemographic background variables are statistically controlled). *But see* Douglas W. Allen et al., *Nontraditional Families and Childhood Progress Through School: A Comment on Rosenfeld*, 50 DEMOGRAPHY 955, 959 (2013) (showing that when those variables are not statistically controlled, school progress of children in households with same-sex couples is significantly different from that of children in other household types).

Using data from the National Longitudinal Study of Adolescent Health ("Add Health"), another research team identified forty-four adolescents parented by female couples who reported they were married or in a marriage-like relationship.[65] Here again, parents' sexual orientation was inferred from their relationship status but was not directly assessed. As a comparison group, the researchers randomly selected other Add Health participants who were being raised by a heterosexual couple and who matched the adolescents from female-couple households on sex, age, ethnic background, adoption status, learning disability status, family income, and parents' educational attainment. The researchers found no significant differences between adolescents from the two different family structures on a large number of key outcome variables, including psychosocial adjustment, school outcomes, romantic relationships and sexual behavior, substance use, delinquency, victimization experiences, and quality of relations with peers.[66] Regardless of family type, having a close relationship with parents and feeling cared for by adults, teachers, and friends were associated with better school adjustment among the adolescents, less delinquent behavior, less substance use, having higher quality peer relations, having more friends in school, and being described by others as more central in their friendship networks.[67]

3. Correlation and Causation

As the previous examples illustrate, many studies in this area compare children raised in one type of family with children from other family structures to assess whether the groups systematically differ in important respects. If differences are observed, they are likely to be interpreted as resulting from the parents' sexual orientation or the child's family structure. That is, the parents' sexual orientation or relationship type is assumed to *cause* the differences in children's levels of adjustment (or whatever variable is the focus of the study).

The meaning of causation and the necessary criteria for demonstrating it have been debated extensively by philosophers of

[65] *See* Wainright & Patterson, *Delinquency*, *supra* note 11, at 527; Wainright & Patterson, *Peer Relations*, *supra* note 11, at 118-19; Wainright, Russell & Patterson, *supra* note 11, at 1889-90.

[66] *See* Wainright & Patterson, *Delinquency*, *supra* note 11, at 528; Wainright & Patterson, *Peer Relations*, *supra* note 11, at 123; Wainright, Russell & Patterson, *supra* note 11, at 1893-95.

[67] *See* Wainright & Patterson, *Delinquency*, *supra* note 11, at 528; Wainright & Patterson, *Peer Relations*, *supra* note 11, at 123; Wainright, Russell & Patterson, *supra* note 11, at 1893-95.

science, and an extended discussion of this topic is beyond the scope of the present paper.[68] For present purposes, let us say that one variable *causes* another when a change in the first variable consistently brings about a subsequent change in the other variable.[69] This definition implies that: (a) the variables must be reliably correlated; (b) the change in the causal variable must temporally precede change in the other variable; and (c) the pattern of change must not be caused by a third variable.[70] I will briefly consider each of these implications in turn.

Consider two variables, labeled X and Y. They are said to be *correlated* if they co-vary, that is, if a change in either variable is consistently accompanied by a change in the other variable. As summarized by the well-known maxim "correlation does not imply causation," such a relationship is necessary but not sufficient for concluding that, for example, X causes Y. Changes in X may indeed cause changes to occur in Y, but the reverse could also be true: Perhaps changes in Y cause X to change. Alternatively, some third variable might cause the pattern of variation observed in both X and Y.

In addition to correlation, another requirement for demonstrating causation is that the cause must temporally precede the effect. Temporal patterns can be readily observed in longitudinal studies (i.e., those that include repeated measurements or observations of the participants over a period of time). In a cross-sectional study (i.e., one that collects data at only a single time-point), the temporal order of events may be impossible to determine. In parenting studies, one might assume that family structure (based on parents' sexual orientation or the gender of the relationship partners) exists prior to the child's current adjustment level. This is indeed the case when the current parents have raised their child together from conception. If, however, the child was adopted or foster-parented, important factors relevant to adjustment may predate the current parent-child relationship. For example, the child may have been exposed to alcohol or drugs *in utero*, or experienced postnatal malnutrition, abuse, or abandonment before being adopted. Thus, before making causal inferences from correlations between family structure and children's well-being, it is important to ensure that the latter variable is not affected by experiences that predate the child's membership in her or his current family. This consideration is

[68] *E.g.*, THOMAS D. COOK & DONALD T. CAMPBELL, QUASI-EXPERIMENTATION: DESIGN AND ANALYSIS ISSUES FOR FIELD SETTINGS 9-36 (1979).

[69] For ease of discussion, "change" is used to refer both to the different quantitative values (e.g., numerical scores on a psychological test) and the different categories (e.g., lesbian parents, heterosexual parents) a variable can assume.

[70] *E.g.*, COOK & CAMPBELL, *supra* note 68, at 9.

especially relevant to the present discussion because lesbian and gay parents appear more likely than heterosexual couples to adopt children from disadvantaged backgrounds or with special needs.[71]

The third requirement for concluding that one variable causes another is to determine that the correlation between them is not due to the influence of a third factor. The ideal method for eliminating this possibility is the controlled experiment, or randomized clinical trial. In the simplest version of this design, the researcher exposes participants to different levels or values of the hypothesized causal variable (usually called the independent variable) and then measures the outcome variable (usually called the dependent variable). The researcher ensures that the experiences of all participants are identical except for the value of the independent variable to which they are exposed. Participants' assignment to a particular value of the independent variable is randomized. When a researcher follows these procedures and observes differences in values of the dependent variable between groups, she or he can conclude that those differences are caused by the independent variable.

Suppose, for example, that a researcher recruits a sample and asks each participant to read a fictitious job applicant's résumé and then watch a video of her job interview. Everyone views the same video and the résumés are identical except for one detail: in one version the list of professional societies to which the applicant belongs includes a gay organization; in the other version it does not. Participants are randomly assigned to read one of the two versions (the independent variable). After they watch the video, participants are asked to rate the applicant on various qualities relevant to the job (e.g., knowledge, dependability) and to indicate whether or not they would recommend that she be hired (these are the dependent variables). If the post-video ratings differ between the two groups, the researcher can conclude that it was due to the independent variable, that is, whether or not the applicant listed the gay organization on her résumé.

Conducting an experiment in which children are randomly assigned to be raised by a heterosexual or same-sex couple obviously is not possible, so researchers use other methods to eliminate the possible effects of extraneous variables in parenting studies. In recruiting the sample, for example, they often ensure that the comparison groups are matched on key variables that are known to affect well-being, such as

[71] *E.g.*, DAVID M. BRODZINSKY, EVAN B. DONALDSON ADOPTION INST., EXPANDING RESOURCES FOR CHILDREN III: RESEARCH-BASED BEST PRACTICES IN ADOPTION BY GAYS AND LESBIANS 16, 40 (2011); GOLDBERG, *supra* note 8, at 66-67; Lavner, Waterman & Peplau, *supra* note 14, at 469.

parents' socioeconomic status or children's age or health status.[72] In addition to or instead of matching, researchers may collect information about potentially relevant variables and use statistical procedures to account for their possible influence on the dependent variable.[73]

When considering variables that may affect the relationships observed between family structure and children's well-being, it is especially important that researchers account for children's differing histories of experiencing family and household instability. Children's psychological and social development is known to be negatively affected by significant disruptions in their early home life (e.g., as a result of parental discord and divorce).[74] Moreover, children raised by lesbian or gay parents are more likely to have experienced such disruptions if they were born into a heterosexual marriage that dissolved when one parent recognized her or his homosexuality. Thus, in at least some samples, children's family type may be correlated with their experience of events such as parental divorce, and the latter are likely to affect children's adjustment. Consequently, variables related to family stability should be accounted for when comparing groups, either through sampling procedures or statistical techniques.

II. ANALYSIS: THREE STUDIES AND THEIR METHODOLOGICAL LIMITATIONS

With this background, the present section evaluates the methodology and findings of three studies that have been cited as evidence that having gay or lesbian parents has negative effects on children. For each study, I summarize the study methodology and key findings, and then

---

72 For examples, see articles cited *supra* note 65.

73 In a 2014 study, for example, researchers found that the adopted children of heterosexual parents exhibited more hyperactivity and conduct problems (based on parents' reports) than the adopted children of gay fathers. Conflating correlation with causation, one might conclude that this pattern demonstrates that having heterosexual parents causes children to have more behavior problems than does having a male-male couple as parents. The researchers also found, however, that heterosexual parents reported experiencing more stress than gay fathers. When variation in stress was statistically controlled, the difference in behavioral problems across family types was not statistically significant. Thus, the differences among the children were more likely due to differing levels of stress related to parenting rather than family structure. *See* Golombok et al., *Adoptive Gay Father Families*, *supra* note 14, 462-63.

74 *See generally* Amato, *supra* note 12 (updating Amato's previous study and finding that children in divorced homes continue to score lower in academics, conduct and social relations); Amato & Keith, *supra* note 12 (conducting a study that found that children in divorced homes score lower on a variety of issues).

demonstrate that the study has serious flaws in one or more of the areas discussed above.

A. The Sarantakos (1996) Study

In 1996, Sotirios Sarantakos published a study of Australian children raised in different types of family constellations.[75] The paper went largely unnoticed by social scientists for several years, probably because it was published in *Children Australia*, an obscure Australian social work journal.[76] Around 2001, it was "discovered" by activists and has since been frequently cited in arguments against marriage equality and adoption rights for sexual minorities.[77]

Sarantakos' nonprobability sample consisted of 174 Australian children of primary school age who were living in households headed by one of three different types of couples: married heterosexual parents, cohabiting heterosexual parents, or cohabiting lesbian and gay parents (n = 58 children in each group).[78] The children were matched on age, sex, and year in school, as well as their parents' occupation, employment status, and levels of education.[79] Based primarily on teachers' subjective evaluations of each child, the three groups were compared on multiple variables related to school and classroom performance, social involvement, and personality characteristics. The specific questions and measures used for these assessments are not detailed in the Sarantakos paper but they apparently were not

75 Sotirios Sarantakos, *Children in Three Contexts: Family, Educational, and Social Development*, 21 CHILD. AUSTL. 23, 23 (1996).

76 Except for a reference to the article made by Sarantakos in one of his own published papers, Sotirios Sarantakos, *Sex and Power in Same-Sex Couples*, 33 AUSTL. J. SOC. ISSUES 17 (1998), it appears not to have been cited in the social science literature until the 2000s. I characterize *Children Australia* as an "obscure" journal because it was not indexed in the major databases for behavioral and social science research (e.g., PsycInfo, PubMed) at the time the article was published. Indeed, obtaining a copy proved somewhat difficult for me because most major research libraries in the United States did not subscribe to it.

77 The study was cited in a 2001 newsletter published by Paul Cameron's Family Research Institute. Its article about the Sarantakos study was headlined "Homosexual Parents: 'Hidden Study' Uncovered!!" FAMILY RESEARCH INST., HOMOSEXUAL PARENTS: 'HIDDEN STUDY' UNCOVERED!! (2001), *available at* https://web.archive.org/web/20010814235549/http://www.familyresearchinst.org/FRR_01_06.html. I have not located earlier citations of the article by authors other than Sarantakos.

78 The letter "n" is used to refer to the number of individuals in the entire sample (usually with a capitalized N) or a subsample (usually with a lower-case n).

79 Sarantakos, *supra* note 75, at 23.

standardized tests or measures. In many cases, teachers simply rated students on a 10-point scale for a particular characteristic.[80]

The study's main findings were that children of homosexual couples were rated significantly lower than other children on most variables: language skills, mathematical abilities, sports, sociability and popularity, attitudes toward school and learning, and parents' school involvement.[81] In addition, "[t]eachers felt that a number of students of homosexual parents were confused about their [gender] identity and what was considered right and expected of them in certain situations."[82]

The study's champions in the United States argued that the observed differences among family types showed that being raised by a same-sex couple has many deleterious effects on children.[83] However, a review of the methodology strongly suggests important third-factor explanations for the differences and highlights an important problem with how some of the variables were operationally defined.

First, all of the children of same-sex couples were born in a previous (presumably heterosexual) relationship "and were subsequently brought into the homosexual relationship."[84] Apparently, none of the children in the same-sex couples group had been raised by the couple since birth and a large number (perhaps most) had experienced parental divorce, many of them in the recent past. The children of married parents apparently had not been subjected to comparable household disruption.[85] Noting the potential relevance of such experiences, Sarantakos acknowledged that the negative effects of divorce and its attendant household disruption on children are well-documented.[86]

---

[80] According to Sarantakos, data were "collected primarily from teachers and only secondarily from parents and children . . . by means of semi-structured questionnaires . . . ." *Id.* at 24. The data were "enriched through telephone interviews" and with "[i]nformation already available through previous studies" of the samples. *Id.* at 24. Unfortunately, the paper does not clearly identify which data were obtained from parents. Apparently, teachers obtained some (perhaps all) of the data provided by the children and subsequently translated their impressions into numerical ratings.

[81] *Id.* at 24-26.

[82] *Id.* at 26.

[83] *E.g.*, Declaration of George A. Rekers in Support of Proposition 22's Motion for Summary Judgment/Summary Adjudication at 9-14, Proposition 22 Legal Def. & Educ. Fund v. City & Cnty. of S.F., No. 04-428794 (Cal. Super. Ct. Dec. 22, 2004); FAMILY RESEARCH INST., *supra* note 77.

[84] Sarantakos, *supra* note 75, at 24.

[85] "Apparently" because Sarantakos does not report the numbers of children in each group who had experienced parental divorce or household disruption. Instead, he offers general characterizations of the children's experiences in the text of the paper. *See id.* at 29-30.

[86] *See generally* Amato, *supra* note 12 (reporting on a meta-analysis of 67 studies on

Thus, an alternative interpretation is that the group differences were caused not by the parents' sexual orientation but by systematic group differences in experiences of family disruption. Because most or all of the children in sexual minority households — but apparently few, if any, of the children with married heterosexual parents — had experienced parental divorce, the effects of family type and experiences of household instability cannot be disentangled.

Second, many of the children of same-sex couples had experienced exceptionally high levels of ostracism and prejudice. According to Sarantakos, "[p]arents and teachers alike reported that comments such as 'the pervs are coming[,]' 'don't mix with the sissies[,]' or 'sisterhood is filthy[,]' made by some pupils, were not uncommon." [87] He reported that high levels of stigma and harassment "were one of the reasons for these children to move to another school, to refuse to go to that school, or even for the parents to move away from that neighbourhood or town."[88]

> In many cases these children have been harassed or ridiculed by their peers for having a homosexual parent, for 'being queer' and even labelled as homosexuals themselves. In certain cases, heterosexual parents advised their children not to associate with children of homosexuals, or gave instructions to the teachers to keep their children as much as possible away from children of homosexual couples. Teachers also reported exceptional cases where a group of 'concerned parents' demanded that three children of homosexuals be removed from their school. Others approached the homosexual parents with the same request. Teachers have reported that children who went through such experiences have suffered significantly in social and emotional terms, but also in terms of scholastic achievement, and have developed negative attitudes to school and learning. These children found it very difficult to adjust in school, to trust friends inside and outside the school, and to join peer groups in general.[89]

Thus, the groups differed systematically on two types of key experiences that are highly likely to affect children's psychological and social adjustment. Experiencing parental divorce and extreme harassment and ostracism — even to the point of being driven out of town or to a different school — were common among the children of

the effects of divorce on children).

[87] Sarantakos, *supra* note 75, at 25.

[88] *Id.*

[89] *Id.* at 26.

same-sex couples but absent among most or all of the children of heterosexual parents.

In addition to these third-factor explanations, the study's operational definition of children's adjustment may have been faulty. Sarantakos cautioned that his reliance on teachers as primary sources of data may have biased the study's outcome, noting that the high levels of sexual stigma and prejudice in the local community were likely to have influenced teachers' subjective ratings of the children. "Obviously, the influence of the attitudes of teachers to life styles on the process of evaluation of students' performance cannot be underestimated."[90] He noted that:

> [Ratings of the children] might have been biased — consciously and/or unconsciously — by the personal views and beliefs of the teachers. In this sense, the attributes of children described in this study might reflect perceptions of attributes rather than actual attributes or differences. Such perceptions might have favoured children of married couples more than children of other couples.[91]

Given these methodological limitations and the dramatic discontinuities between the Sarantakos findings and other research in this area, the alternative interpretation of the results is highly plausible. Although a correlation was observed between parent sexual orientation and children's adjustment and school performance, the study design does not permit conclusions to be drawn about a causal relationship. Indeed, rather than being due to parents' sexual orientation, the observed differences among children appear to be more likely a result of the confounding of their family type with their history of household disruption and exposure to an extremely hostile environment. The high levels of stigma and prejudice that characterized the research setting may also have influenced the teachers' ratings of the children.[92]

B. The Regnerus (2012) Study

A second study whose findings differed starkly from those of other research in this area was published in 2012 by sociologist Mark Regnerus.[93] Labeled the New Family Structures Study ("NFSS"), the data were obtained from a sample (N = 2988; age = 18–39 years) drawn

[90] *Id.*

[91] *Id.* at 30.

[92] *Id.*

[93] For details, see Regnerus, *How Different*, *supra* note 3.

from a panel assembled by Knowledge Networks ("KN").[94] Before discussing the NFSS, it is necessary to briefly describe the KN panel, which consists of more than 50,000 people. Unlike most Internet surveys, the panel constitutes a true national probability sample, as do subsamples derived from it. The sample was initially created using the same methods that are employed for obtaining nationally representative samples for traditional telephone and household surveys, as described earlier. Once contacted, sample members were invited to join the panel with the understanding that they would be asked to complete questionnaires a few times each month, and were free to decline to participate in any given survey. In contrast to most online surveys, the panel is not restricted to individuals with Internet access. Sample members who lack access are provided with it free of charge in return for their participation, along with the necessary computer hardware. Using these methods, Knowledge Networks has been able to create a large, nationally representative panel that includes people who would otherwise be excluded from online studies. Because they have collected extensive background data from each panel member, they are able to create subsamples that represent specific demographic and social subgroups of the population.[95]

---

[94] *Id.* at 755-56. Knowledge Networks is now called GfK.

[95] For more information about the panel, see generally *GfK KnowledgePanel*, GFK, http://www.gfk.com/us/Solutions/consumer-panels/Pages/GfK-KnowledgePanel.aspx (last visited Nov. 11, 2014). Although some critics of the Regnerus study have questioned the sampling methodology, the KN/GfK procedures are considered by many public opinion researchers to be basically sound and were found in one study to yield data of better quality than traditional telephone samples. *Compare* Darren E. Sherkat, *The Editorial Process and Politicized Scholarship: Monday Morning Editorial Quarterbacking and a Call for Scientific Vigilance*, 41 SOC. SCI. RES. 1346, 1348-49 (2012) (questioning the representativeness of the KN panel), *with* LinChiat Chang & Jon A. Krosnick, *National Surveys via RDD Telephone Interviewing Versus the Internet: Comparing Sample Representativeness and Response Quality*, 73 PUB. OPINION Q. 641, 675 (2009) (finding that the KN panel yielded a national sample whose representativeness was comparable to that obtained through traditional telephone sampling methods, and responses that were more accurate and of higher quality, based on a variey of criteria). I myself have utilized KN panel data to collect data from U.S. national probability samples of the general adult population and self-identified lesbian, gay, and bisexual adults and have described the procedures and response rates in some detail. *See* Gregory M. Herek, *Hate Crimes and Stigma-Related Experiences Among Sexual Minority Adults in the United States: Prevalence Estimates From a National Probability Sample*, 24 J. INTERPERSONAL VIOLENCE 54, 58 (2009); Herek et al., *supra* note 58, at 179-80; Aaron T. Norton & Gregory M. Herek, *Heterosexuals' Attitudes Toward Transgender People: Findings from a National Probability Sample of U.S. Adults*, 68 SEX ROLES 738, 742 (2013). This is not to deny the KN/GfK panel's potential limitations. For example, people who complete questionnaires on a regular basis as part of a panel may respond differently than a sample of novice participants. To date, the research that has addressed this issue

For the NFSS, Regnerus assigned respondents to different categories depending on their family type prior to age eighteen. Because of their centrality to the validity of the conclusions he drew from the data, his operational definitions and categorization procedures warrant a somewhat detailed description. At the beginning of the questionnaire, all respondents were asked Screening Question #S2, "Did you live together with BOTH your biological mother AND biological father the entire time from when you were born until age 18 (or until you left home to be on your own)?" Respondents who answered "yes" to #S2 were assigned to the "IBF" (intact biological family) category ($n = 919$), and were not asked any of the subsequent screening questions.[96]

Respondents who answered "no" to #S2 were asked several follow-up screening questions. One of these (Question #S7) was "From when you were born until age 18 (or until you left home to be on your own), did either of your parents ever have a romantic relationship with someone of the same sex?" Respondents were placed in a category labeled "GF" (child of a gay father) if they responded "No" to #S2 and "Yes, my father had a romantic relationship with another man" to #S7. They were assigned to a category labeled "LM" (child of a lesbian mother) if they answered "no" to #S2, did *not* report on #S7 that they believed their father had a romantic relationship with another man, and replied to #S7 with "Yes, my mother had a romantic relationship with another woman."[97]

Those assigned to the GF ($n = 73$) and LM ($n = 163$) categories were then asked one or both of the following questions: "Did you ever live with your mother while she was in a romantic relationship with another woman?" and "Did you ever live with your father while he was in a romantic relationship with another man?"[98] Respondents assigned to

suggests that the response patterns of the two groups probably do not differ in substance, but novice respondents may be more likely to respond "don't know" to some types of attitude questions. Johannes Binswanger et al., *Panel Conditioning in Difficult Attitudinal Questions*, 77 PUB. OP. Q. 783, 790-91 (2013); Vera Toepoel et al., *Effects of Design in Web Surveys: Comparing Trained and Fresh Respondents*, 72 PUB. OP. Q. 985, 1002-03 (2008). To the extent that any such differences exist, there is no reason to believe they would introduce any bias specifically related to the present discussion.

96 Regnerus, *How Different*, *supra* note 3, at 757; Mark D. Regnerus, *NFSS Survey Questionnaire*, UNIV. TEX. AUSTIN 1, 2 (2012), http://www.prc.utexas.edu/nfss/documents/NFSS-Survey-Instrument.pdf [hereinafter *NFSS Survey*].

97 If respondents believed that both parents had a same-sex romantic relationship they were assigned to the GF group. According to the study, this procedure was used for data analytic purposes to maximize the size of the GF group, which "is the smallest and most difficult to locate randomly in the population." Regnerus, *How Different*, *supra* note 3, at 758.

98 Question wording obtained from the article, *id.* at 756, and the NFSS

the IBF category apparently were not asked their beliefs about their parents' same-sex romantic relationships.[99]

Other screening questions asked whether the respondent lived with one or more adoptive parents — including a biological relative or step-parent — or a single biological parent (with or without a cohabiting partner of the other sex).[100] A question located later in the survey (#Q4) asked if the respondent's biological parents were currently married, if they had been married but later separated or divorced, if they had been married but one or both was now deceased, if they had cohabited without marrying, or if they had never cohabited or married.[101] Regardless of their answers to these questions, all respondents who reported believing that one or both parents had a same-sex romantic relationship were placed in the GF or LM group.[102] Consequently, these two groups were heterogeneous with respect to their members' experiences with various kinds of household transitions while growing up. By contrast, respondents who did not report believing that a parent had a same-sex romantic relationship were categorized into more homogeneous groups, with the IBF group consisting exclusively of individuals who had never experienced the loss of either biological parent through divorce, separation, death, or abandonment.[103]

When Regnerus compared the GF and LM respondents to the IBF respondents, he found numerous differences. For example, the GF and LM respondents reported significantly lower educational attainment, more depression, lower quality in their current romantic relationship, more frequent smoking, more arrests and guilty pleas to non-minor offenses, more male sex partners, and (among women but not men respondents) more female sex partners.[104] According to Regnerus, his data showed that children whose mothers or fathers have ever been in a same-sex relationship grow up to have many more psychological,

questionnaire. Regnerus, *NFSS Survey*, *supra* note 96, at 3-4.

[99] I say "apparently" because the Regnerus paper is not explicit about this point. However, according to the questionnaire, participants who responded yes to question #S2 were taken directly to the main questionnaire section (#Q2) without being asked any of the follow-up screening questions. This aspect of the classification system ignores the fact that it is possible to grow up living with both of one's biological parents (i.e., answer "yes" to #S2) while also believing that one or both of them had same-sex romantic relationships (i.e., answer "yes" to #S7). *See* Regnerus, *NFSS Survey*, *supra* note 96, at 1, 3.

[100] *Id.* at 2-3.

[101] *Id.* at 6.

[102] Regnerus, *How Different*, *supra* note 3, at 757.

[103] *Id.* at 757-58.

[104] *Id.* at 763-64.

social, and behavioral problems than children raised by their married biological mother and father.[105] These findings were heralded by many conservative activists as demonstrating that previous studies showing no relationship between parent sexual orientation and children's outcomes were invalid. [106]

The conclusions drawn by Regnerus, however, are untenable because of the fundamental limitations in the study sample and in his operational definitions of the key categorization variable. In his paper and in public comments, Regnerus contrasted his findings to those of studies that compared children raised by heterosexual versus same-sex couples, or heterosexual versus sexual minority single parents.[107] Many of his conclusions involved generalizing from respondents in the GF and LM groups in his sample to the general population of young adults raised by single or coupled lesbian, gay, or bisexual parents. Thus, the study was framed as a comparison of adult outcomes for children raised by heterosexual versus sexual minority parents.

Because of the procedures Regnerus used for operationally defining family type, however, a serious discrepancy exists between his

---

[105] *Id.* at 766.

[106] The publication of the Regnerus paper in *Social Science Research* was greeted with considerable controversy. A group of more than 100 social scientists and mental health professionals signed a letter to the journal criticizing the study. *See* Gary J. Gates et al., *Letter to the Editors and Advisory Editors of* Social Science Research, 41 SOC. SCI. RES. 1350 (2012). The editor of *Social Science Research* commissioned an audit of the editorial process through which the paper was accepted, which concluded that the journal's peer-review procedures had been inadequate. Sherkat, *supra* note 95, at 1348-49. Regnerus was subjected to an investigation by his university, which concluded he was not guilty of scientific misconduct. Memorandum from Robert A. Peterson for Executive Vice-President and Provost S. Leslie 1, 3 (August 24, 2012) (on file with The University of Texas at Austin), http://www.utexas.edu/opa/wordpress/news/files/Regnerus-Inquiry-Report.pdf. Meanwhile, opponents of marriage equality defended the study and Regnerus's conclusions, as did some academics and researchers. *See* Jeremy Weber, *Social Scientists Defend Mark Regnerus' Controversial Study on Same-Sex Parenting*, CHRISTIANITY TODAY (July 10, 2012), http://www.christianitytoday.com/gleanings/2012/july/social-scientists-defend-mark-regnerus-controversial-study.html. A discussion of these events is beyond the scope of this article. For a more detailed account, see, for example, Tom Bartlett, *Controversial Gay-Parenting Study Is Severely Flawed, Journal's Audit Finds*, THE CHRONICLE OF HIGHER EDUC. (July 26, 2012), http://chronicle.com/blogs/percolator/controversial-gay-parenting-study-is-severely-flawed-journals-audit-finds/30255; Eric Eckholm, *Opponents of Same-Sex Marriage Take Bad-for-Children Argument to Court*, N.Y. TIMES, Feb. 23, 2014, at A16, *available at* www.nytimes.com/2014/02/23/us/opponents-of-same-sex-marriage-take-bad-for-children-argument-to-court.html. *See generally* Plaintiffs' Motion in Limine to Exclude Testimony of Mark Regnerus at 7-19, Deboer v. Snyder, 973 F. Supp. 2d 757 (E.D. Mich 2014) (No. 12-cv-10285) (challenging Regnerus's qualifications as an expert witness in a suit to overturn Michigan's ban on marriage for same-sex couples).

[107] Regnerus, *How Different*, *supra* note 3, at 766.

conceptual and operational definitions of participants' family background. As a result, the final sample and the study's target population were clearly mismatched. Regnerus did not obtain data directly from the participants' parents about their sexual orientation, sexual behavior history, past romantic involvements, or any other variables. Nor were respondents asked directly about their beliefs concerning their parents' sexual orientation or past sexual activity. Instead, the key criterion for placing respondents into the GF and LM groups was their *belief* that a parent ever had a *romantic* relationship with a person of the same sex.[108] The extent to which the belief was accurate and how respondents interpreted the term "romantic" are not known. Thus, the GF and LM groups are most accurately described as consisting of young adults who believed that one or both of their biological parents ever had a "romantic" relationship with someone of the same sex.

Regnerus contributed to the misapprehension that his study was about lesbian and gay parenting by framing it as a corrective to prior research on children parented by lesbians or gay men, by using the terms "lesbian" and "gay" interchangeably to refer both to sexual orientation identities and to same-sex sexual behavior, and by mislabeling his key comparison groups with the initials "LM" (child of a lesbian mother) and "GF" (child of a gay father).[109] From his operational definitions, however, it is impossible to determine how many of the NFSS respondents had parents who could be considered lesbian, gay, or bisexual. Clearly, few participants were raised in a household headed by a same-sex couple. According to the initial Regnerus paper, 77% of the participants in the "Lesbian Mothers" group and 98% in the "Gay Fathers" group had *not* lived with the parent and her or his same-sex partner for at least 3 years.[110] In fact, of the 163 respondents in the LM group, Regnerus noted that "only 19 [i.e., about 12%] spent at least five consecutive years [living with their mother and her partner], and six cases [4%] spent 10 or more consecutive years together."[111] Only two "reported living with their mother and her [female] partner uninterrupted from age 1 to 18."[112] Among the 73 respondents categorized as living with a gay father, fewer than 2% had

---

[108] *Id.* at 755-56.

[109] *Id.* at 757-58.

[110] *Id.* at 757.

[111] Mark Regnerus, *Parental Same-sex Relationships, Family Instability, and Subsequent Life Outcomes for Adult Children: Answering Critics of the New Family Structures Study with Additional Analyses*, 41 SOC. SCI. RES. 1367, 1370 (2012).

[112] *Id.*

lived with their father and his male partner for even three years.[113] The extremely small number of respondents who resided for any length of time with a parent in a same-sex couple precludes any meaningful statistical comparisons between them and other study participants.

As noted above, respondents who were placed in the LM and GF groups were highly heterogeneous in family backgrounds and experiences (e.g., divorce, loss of a parent, foster care). Indeed, in an article written for *Slate*, Regnerus stated that respondents in these groups grew up with less household stability than any of the others in his study.[114] Nevertheless, he compared them to respondents who were selected specifically because they did *not* experience divorce or other family disruptions while growing up. Thus, rather than providing insights into outcomes for adults who were raised by one or two lesbian, gay, or bisexual parents, the study results are more plausibly interpreted as echoing previous findings that experiencing divorce and family disruption during childhood is associated with problems in adulthood.

C. The Allen (2013) Study

A third study purporting to demonstrate significant differences among children raised in different family types was conducted by economist Douglas Allen using 2006 Canadian Census data.[115] He compared high school graduation rates across family structures among young adults (ages 17–22) currently residing with a parent.[116] He concluded that those living in a gay or lesbian family are less likely to graduate than their counterparts living in a family headed by a heterosexual married couple.[117]

---

[113] Regnerus, *How Different*, *supra* note 3, at 757. Although Regnerus reported this finding as a percentage, it is worth noting that two percent of 73 equals 1.46. Thus, it is most likely the case that this "group" consisted of only one (1) respondent.

[114] Mark D. Regnerus, *Queers as Folk: Does It Really Make No Difference If Your Parents Are Straight or Gay?*, SLATE (June 11. 2012, 6:02 AM), http://www.slate.com/articles/double_x/doublex/2012/06/gay_parents_are_they_really_no_different_.html.

[115] Allen, *High School Graduation Rates*, *supra* note 3, at 653.

[116] In his original article, Allen stated that children of same-sex couples were "those who respond affirmative to the question: 'Are you a child of a male (female) same-sex married or common law couple?'" *Id.* at 643. However, the census form did not include any such question. In a published erratum, Allen later acknowledged this, explaining that these children were actually "identified by a Statistics Canada created variable 'RELATIONSHIP TO PERSON 1' which included 'CHILD OF A MALE (FEMALE) SAME-SEX MARRIED OR COMMON LAW COUPLE' among the possible categories for that variable." Douglas W. Allen, *Erratum to: High School Graduation Rates Among Children of Same-Sex Households*, 12 REV. ECON. HOUSEHOLD 207, 207 (2014).

[117] Allen, *High School Graduation Rates*, *supra* note 3, at 653.

There are several problems with Allen's conclusion, including that the data analyses highlighted in the body of the paper only partially support it, and other analyses reported in an appendix do not support it at all. As noted above, experiencing household disruption is often a source of problems for children, regardless of their parents' sexual orientation. Lacking other indicators of young adults' history of experiencing household disruption while growing up, Allen used changes in residence as a proxy. His report emphasizes an analysis that controlled statistically for residence change during the previous year, that is, when the young adults were 16–21 years old.[118] It also controlled for numerous other variables, including parents' current marital status.[119] This analysis revealed significantly lower rates of high school graduation among young adults who were currently living with parents in a male-male couple, compared to those living with heterosexual married parents, but no statistically significant difference for the children of female couples.[120]

During the one-year period considered in that analysis, 19% of the lesbian households had relocated, compared to 8% of the gay male households and 7% of the heterosexual married households.[121] When the time frame is expanded to include the previous five years, however, the group differences are more dramatic: 39% of young adults residing in a gay male household and 60% of those residing in a lesbian household had changed residence in the previous five years, compared to 24% of those in heterosexual married households.[122]

When Allen repeated his earlier analysis using household mobility during the previous five years (rather than one year) as a statistical control, graduation rates were not significantly different across household types (i.e., headed by male couples, female couples, or married heterosexual couples).[123] This key finding was not reported in the body of the paper but instead appeared without accompanying discussion in an Appendix comprising Table 8.[124] The only mention of

118 Table 5 provides the regression data. *Id.* at 645-49 tbl.5.

119 As reported in Table 4, fewer than half of the same-sex couples (45% of males and 20% of females) were legally married at the time of data collection. *Id.* at 647 tbl.4. The parents in the married heterosexual comparison group were all, by definition, married. *Id.*

120 *See id.* at 648 & tbl.5 (Model 3).

121 *Id.* at 647 tbl.4.

122 *Id.*

123 *Id.* at 654 tbl.8.

124 *Id.* (Model 3)

Table 8 in the body of the paper appeared on pages 645 to 646, where Allen stated:

> All of the regressions in [Table 5] control for whether or not the family moved . . . [within] the past year. Table 8 in the appendix reports on another three logit regressions with the same dependent variable and the same right hand side variables, except for the variable used to control for family mobility — it uses the mobility measure "did child move within past 5 years." There is no qualitative difference in the estimates when using the different mobility controls.[125]

Table 8 shows that, with all variables included in the analysis, the differences in odds ratios for high school graduation between young adults residing with their married heterosexual parents and their counterparts living with parents in a married or common-law same-sex couple are not significantly different.[126] Despite Allen's assertion to the contrary, this loss of statistical significance is an important "qualitative difference" between the analyses.

Even if the graduation rates had differed significantly across family structures, however, there are two additional reasons why meaningful conclusions about a relationship between family structure and children's educational outcomes could not be drawn from the data. First, although data were available concerning household change during the five years prior to data collection (i.e., when the children were 12 to 17 years old), the Canadian Census is not a longitudinal study and thus does not include data about parents' prior marital history or the child's earlier family history.[127] Thus, it is not possible to control for household disruption earlier in the children's life or for parents' earlier divorce, separation, or loss of a spouse. As Allen acknowledged, "*this paper does not study the effect of growing up in a same-sex household*,

[125] *Id.* at 645-46 (footnote omitted). Allen discusses two types of controls for mobility. *See id.* at 646 n.34. The first is whether the child's residence had moved within the previous five years. *Id.* The second is whether the child had changed census metropolitan areas over the past 5 years. *Id.* Apparently, equations were constructed using each type of mobility measure, but which type was used as a control variable in the analyses reported in Tables 5 and 8 is not clear.

[126] This does not mean that the proportions of stay-at-home high school graduates in each household type changed when new variables were included in the analysis. Rather, it indicates that the observed differences in graduation rates were substantially due to the other variables in the analysis, especially household relocation and parents' marital status.

[127] Allen, *supra* note 3, at 654.

but rather examines the association of school performance for those children who lived with same-sex parents in 2006."[128]

Second, the data that Allen analyzed excluded 17- to 22-year-olds who no longer resided with their parents at the time of the Census.[129] This is a noteworthy omission. After successfully completing high school, many young people leave their parents' home to join the work force, attend college, enlist in the military, or pursue other paths. It cannot be assumed that the family structure patterns of this key group match those of the stay-at-home group studied by Allen. It could be hypothesized, for example, that children raised in lesbian and gay male households are more likely to leave home and live independently after they graduate from high school, perhaps because they are more mature and self-reliant than those raised in married heterosexual households. Without data from the young adults who no longer lived with their parents, this hypothesis cannot be tested. In short, whereas the appropriate target population for Allen's research question comprises all Canadian 17- to 22-year olds, the sample systematically excludes a key segment of that population. The findings from Allen's sample do not illuminate the relationship between family type and educational achievement in the larger population.

CONCLUSION

Whereas the studies conducted by Sarantakos, Regnerus, and Allen have been cited as evidence that children are negatively affected by having parents who are lesbian, gay, bisexual, or in a same-sex couple, I have demonstrated that each study suffers from one or more serious methodological flaws.

Sarantakos's finding that children of lesbian and gay parents showed markedly worse adjustment and school performance than children of heterosexual parents is most likely a consequence of the study's conflation of parents' sexual orientation with children's experiences of divorce, family disruption, and an extremely hostile social climate.

The Regnerus data do not permit conclusions to be drawn about children raised by same-sex couples because, as he acknowledged, the sample included only two individuals who had lived with a same-sex parent couple since infancy. Because of the operational definitions Regnerus used for identifying participants with a "lesbian mother" or "gay father," his sample cannot even be confidently characterized as consisting of young adults raised by a lesbian or gay parent.

---

128 *Id.* at 639 n.9 (emphasis added).

129 *Id.* at 643 ("[A]ll children *living with* a parent within the home were selected.").

The Allen study actually demonstrated that, once parental marital status and household changes during the previous five years are controlled, observed differences in high school graduation rates between stay-at-home young adults in sexual minority and heterosexual married households are not statistically significant. Because his data set excludes a key sector of the target population — namely, 17- to 22-year-olds who were no longer residing with their parents when the Canadian Census was conducted — its usefulness for assessing relationships between family structure and children's graduation rates is extremely limited.

I conclude that these three studies have little or no relevance to empirically-based discussions of parenting and sexual orientation. The current social science research literature can be accurately characterized as failing to find reliable differences in psychological and social adjustment between children with heterosexual parents and those who have sexual minority parents or were raised by same-sex couples.

Through my discussion of the Sarantakos, Regnerus, and Allen studies, I have tried to demonstrate that, as in other areas of social science research, evaluating the contributions made by any empirical study of family structure and children's well-being requires critical consideration of methodological issues. It also requires examination of the study's findings in light of the patterns observed across other studies addressing similar questions. Each of those studies will have its own limitations — none are perfect. It is important to recognize that empirical research can have methodological shortcomings but still contribute to scientific knowledge. Indeed, regardless of the topic, most published reports of peer-reviewed social science research include discussion of the study's limitations and how future research might improve upon its design. Thus, I do not intend to suggest that empirical research must be methodologically flawless to yield useful information. The methodology of some studies, however, is simply inadequate for addressing the research question they purport to investigate. This is the case for the Sarantakos, Regnerus, and Allen studies.

My discussion here has focused mainly on methodological issues, and I have tried to provide information that will be useful to readers who are not themselves social scientists when they evaluate new research in this and related areas. However, focusing exclusively on the research methods and findings can distract us from the task of critically questioning the cultural backdrop for research in this area. Sexual stigma fosters a differences-are-deficits assumption: if differences were to be observed among children reared in different family structures, they would be assumed to indicate a problem on the part of sexual minority parents and same-sex couples. Moreover, stigma creates the further assumption that such differences, if they were found to exist, would justify denying same-sex couples the right to marry and denying their children the right to have their relationship with their nonbiological parents legally recognized.

Comparable assumptions are not typically voiced when differences in childrearing outcomes are observed across groups defined by characteristics other than sexual orientation, such as socioeconomic status, race, ethnicity, and geographic area of residence. If children in a category defined by one of those variables are observed to manifest more adjustment problems or lower scholastic achievement than their peers, society's response typically is to seek ways to strengthen those families and help the children, not to deny adults in that category the legal right to marry or parent. With sexual minorities, however, the belief that such differences exist (even in the absence of empirical support for that belief) is perceived by some as sufficient justification for punishing parents and children by denying them the benefits that society bestows on heterosexuals and their families. When these beliefs are enshrined in law and policy, the resulting differential treatment is an instance of structural stigma.

Many thousands of children are already being raised by parents in same-sex relationships and by single lesbian, gay, and bisexual adults.[130] Thus, whether sexual minority adults should raise children is not really an issue for debate. Instead, the important question is whether the children of same-sex couples benefit from or are harmed by laws that prevent their parents from marrying or that recognize the child's relationship to only one of her or his parents. As noted above, the factors that facilitate children's adjustment are well established: positive relationships with parents or caretakers; positive and supportive relationships between those adults, which create a stable household; and a home environment that affords adequate economic, social, and physical resources while shielding children from poverty and social isolation.[131] These resources are more likely to be available when children's relationships with their parents and their parents' relationship with each other are legally recognized. As the U.S. Court of Appeals for the Ninth Circuit observed in its ruling striking down Idaho and Nevada laws barring same-sex couples from marrying, "[r]aising

[130] *See generally* GARY J. GATES, LGB FAMILIES AND RELATIONSHIPS: ANALYSES OF THE 2013 NATIONAL HEALTH INTERVIEW SURVEY (2014) (estimating that nearly 200,000 children under the age of 18 are currently being raised by same-sex couples, with more than one million other children being raised by single lesbian, gay, or bisexual adults).

[131] *See supra* notes 10–11 and accompanying text.

children is hard; marriage supports same-sex couples in parenting their children, just as it does opposite-sex couples."[132] Thus, assuring that the law permits them to receive such support serves the best interests of lesbian, gay, and bisexual parents and their children.

132 Latta v. Otter, No. 12 17666, 2014 WL 4977682, at *6 (9th Cir. Oct. 7, 2014).