Arch Sex Behav (2014) 43:191–196
DOI 10.1007/s10508-013-0239-1

INVITED COMMENTARY

# Considering Alternative Explanations for the Associations Among Childhood Adversity, Childhood Abuse, and Adult Sexual Orientation: Reply to Bailey and Bailey (2013) and Rind (2013)

**Andrea L. Roberts · M. Maria Glymour · Karestan C. Koenen**

Published online: 24 December 2013


Using a nationally representative U.S. dataset, we noted the established association between child abuse and same-sex sexuality and asked whether this association was most likely due to children's sexual orientation influencing risk of abuse, as commonly assumed, or whether child abuse might affect sexual orientation (Roberts, Glymour, & Koenen, 2013). We hypothesized that abuse influenced orientation and used an instrumental variable approach to assess this hypothesis. Specifically, because childhood adversities are known to influence risk of abuse, but have no known direct influence on sexual orientation, we hypothesized that, if abuse affects sexual orientation, adversities that increase the risk of abuse should also predict higher prevalence of same-sex sexual orientation.

We found support for this hypothesis in that childhood adversity predicted childhood sexual abuse; that childhood adversity also predicted same-sex sexual attraction, partners, and identity; and that childhood adversity was independent of same-sex sexual attraction, partners, and identity when accounting for childhood abuse. Using instrumental variable models, we estimated that half to all of the elevated risk of childhood abuse among persons with same-sex sexuality compared to heterosexuals was due to the effects of abuse on sexuality. Since the publication of our article, a new study using different data found that gay men, lesbians, and bisexual persons compared with heterosexuals were more likely to experience household-level adverse circumstances in childhood, including household mental illness, household substance abuse, an incarcerated household member, and (for bisexuals only) parental separation or divorce (Andersen & Blosnich, 2013). These findings again raise the question of what might account for the higher prevalence of household-level childhood adversities that are risk factors for childhood abuse among families of sexual orientation minorities.

We appreciate the thoughtful commentaries from Bailey and Bailey (2013) and Rind (2013) and thank the Editor for the opportunity to respond. Our article addressed a sensitive issue. Persons who identify as gay, lesbian, or bisexual have been and continue to be discriminated against both individually and institutionally. Homosexuality was a diagnosable mental disorder as recently as DSM-II. Because of this, even to ask the question of what factors contribute to sexual orientation is sensitive. Rind takes our research to imply that homosexual orientation is "abnormal," "pathological" or "maladaptive." We do not state this and we strongly do not believe it. Our research was conducted in the spirit of investigating individual differences in human behavior as is done with traits such as personality. We disagree with those who would apply our findings for political goals that would harm or demean persons who identify as gay, lesbian, or bisexual. However, we do not believe the fear that someone might misuse or misinterpret our findings should preclude research on the origins of sexual orientation or on the link between sexual orientation and childhood abuse.

The instrumental variable models cannot be proven; they are interpretable as causal only with additional causal assumptions. We contrast here the assumptions required for our

A. L. Roberts (✉)
Department of Social and Behavioral Sciences, School of Public Health, Harvard University, Kresge Building, 677 Huntington Ave., Boston, MA 02115, USA
e-mail: aroberts@hsph.harvard.edu

M. M. Glymour
Department of Epidemiology and Biostatistics, University of California San Francisco School of Medicine, San Francisco, CA, USA

K. C. Koenen
Mailman School of Public Health, Columbia University, New York, NY, USA

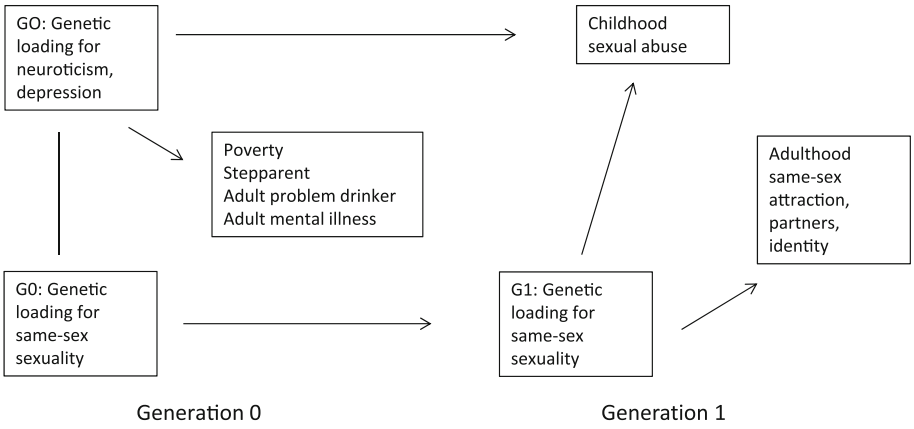


**Fig. 1** Bailey and Bailey: genetics as a common cause of instruments, childhood abuse and, same-sex sexuality

interpretation with the assumptions and implications of the alternative proposals from Bailey and Bailey (2013) and Rind (2013).

Bailey and Bailey proposed that same-sex sexuality is influenced by a genetic factor that also predicts parental difficulties, such as divorce, mental illness, poverty, and drug use. They proposed genetic factors that increase the risk for neuroticism as one such possibility. Under this hypothesis, the association between, for example, presence of stepparents in early childhood and same-sex behavior is due to confounding by the gene (Fig. 1). We note that Bailey and Bailey's hypothesis implies that gay men and lesbians carry genes—passed down from their parents—that increase their risk of mental illness, alcohol use, poverty, and instability in long-term relationships. To our knowledge, there is no genetic research that supports this possibility.

To investigate the likelihood that the causal structures proposed by Bailey and Bailey could account for the associations present in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) data, we conducted several simulations. Our objective was to simulate a world in which the statistical associations in the data could arise from the causal structure proposed by Bailey and Bailey, to assess whether this structure was plausible (for details of the simulations and code, see Appendix). These simulations indicate that the causal structure proposed by Bailey and Bailey (Fig. 1) can create the association between stepparents and same-sex identity found in NESARC only if very strong genetic effects on these phenotypes exist. For example, to fulfill Bailey and Bailey's hypothesis, the risk allele must account for approximately 14 % of the mother's neuroticism and 15 % of the child's probability of having a same-sex identity. These are stronger, by an order of magnitude, than any established genetic determinant for any mental health or complex behavioral outcome. For example, a polygenic risk score for schizophrenia comprised of more than 37,000 single-nucleotide polymorphisms (SNPs) explained at most 3 % of the risk of schizophrenia (Purcell et al., 2009). Even if there were a genetic determinant that explained 14 % of maternal neuroticism, to generate the associations present in the NESARC data, we assumed the neuroticism risk allele had the same effect size on likelihood of same-sex identity as on neuroticism. This seems unlikely given prior evidence on shared heritability of complex phenotypes in the same domain (Purcell et al., 2009). Even assuming these strong genetic effects, we were only able to obtain the association between having a stepparent before age 5 and same-sex identity found in NESARC if mother's neuroticism accounted for 50 % of the likelihood of having a stepparent.

In sum, we simulated data under a range of assumptions and were unable to generate any data set that was consistent with the causal structure proposed by Bailey and Bailey, current knowledge of genetic determinants of psychological and behavioral traits, and the observed statistical patterns in the NESARC data. We therefore conclude that their proposed causal structure is extremely unlikely. In our simulations, we considered many possible alternatives, but we inevitably did not explore the complete universe of possible models and made assumptions about the functional form of the causal links (e.g., linear effects). We therefore cannot rule out that there is some alternative, complex data-generating mechanism that would be consistent with both the proposed causal structure and the observed data, and we invite Bailey and Bailey to propose such a mechanism.

We now turn to Rind's (2013) hypothesized causal structure. Rind suggests that the childhood adversities we examined (poverty, parental alcohol problem, parental mental illness, and having a stepparent) "weaken normative controls," which leads to increased likelihood of acknowledging or acting on existing same-sex attractions. It is unclear why Rind does not allow that experiences of child maltreatment may be powerfully non-normative in themselves. We stated this possibility in our article:

> …abuse survivors may feel stigmatized and different from others and may, therefore, be more willing to behave in ways that are socially stigmatized, including acknowledging same-sex attraction or having same-sex

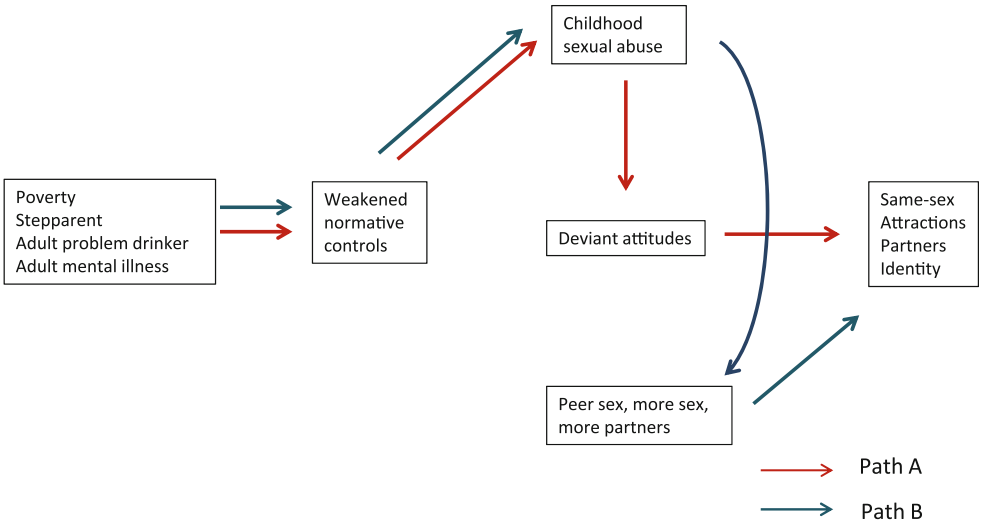


**Fig. 2** Rind's pathways from instruments through childhood abuse to same-sexuality

**Table 1** Prevalence of same-sex attraction, partners, and identity by childhood circumstances among men and women not exposed to childhood abuse, NESARC ($n = 10{,}375$)

| | *N* | Same-sex attraction | Same-sex partners | Same-sex identity |
|---|---|---|---|---|
| | | | % | |
| Poverty | | | | |
| No | 9,371 | 5.3 | 2.3 | 1.0 |
| Yes | 960 | 5.0 | 2.3 | 1.0 |
| Parent alcohol problem | | | | |
| No | 9,369 | 5.3 | 2.3 | 1.0 |
| Yes | 1,006 | 4.6 | 2.2 | 1.0 |
| Stepparent before age 5 | | | | |
| No | 9,477 | 5.3 | 2.4 | 1.1 |
| Yes | 182 | 1.7 | 1.7 | 0.0 |
| Parent mental illness | | | | |
| No | 10,051 | 5.3 | 2.3 | 1.0 |
| Yes | 326 | 4.9 | 3.1 | 1.0 |

*Note* Ns differ slightly for each childhood circumstance/sexuality combination due to missing responses

> partners (Saewyc et al., 2006)…. It would also follow that in societies where same-sex sexuality is more accepted and less stigmatized, prevalence of same-sex sexual orientation would be higher and sexual orientation disparities in abuse would be lower. (p. 169)

If we replace "socially stigmatized" with "counternormative," the argument is the same. In fact, Rind's causal diagram indicates several pathways by which childhood maltreatment affects sexual orientation (we highlight two of these pathways in Fig. 2).

It is also possible to test Rind's hypothesis using the NESARC data. Were Rind's proposed causal structure accurate, non-normative childhood experiences would be associated with same-sex sexuality regardless of child abuse status. We therefore examined the association of our instruments with same-sex sexuality among persons who did not experience childhood abuse. Table 1 shows the prevalence of same-sex sexuality by childhood adversity among men and women who did not experience childhood abuse. Among persons reporting no abuse, the prevalence of same-sex attraction, partners, and identity was, in general, the same or lower in those who experienced poverty, parent alcohol problem, a stepparent or parental mental illness compared with those who did not. Although not conclusive, these data suggest that there is no effect of these non-normative experiences on sexuality except when child abuse occurs.

Bailey and Bailey incorrectly asserted that we rejected the possibility that nascent childhood sexual orientation affects both childhood maltreatment and adult sexual orientation because the instruments (childhood adversity) were correlated with adult minority sexual orientation. On the contrary, we reject this possibility because the instruments were uncorelated with adult sexual orientation when conditioning on childhood maltreatment. If childhood adversity directly affected nascent childhood sexual orientation, which affected both maltreatment and adult orientation, the correlation between childhood adversity and adult orientation should not be eliminated by adjustment for maltreatment. We appreciate that Bailey and Bailey focused on the key assumptions for our instrumental variable models: (1) there are no unmeasured causes of childhood adversity (the instrumental variables) and sexual orientation; and (2) childhood adversity does not affect sexual orientation via some other mechanism, unrelated to childhood abuse. They argue that these assumptions may not be true and proposed an alternative explanation for the observed empirical patterns. Although we agree that the assumptions may not be true, the specific alternative proposed by Bailey and Bailey seems implausible. We welcome additional theorizing on plausible alternatives and believe it will advance our understanding of both childhood maltreatment and the origins of sexual orientation.

In conclusion, although instrumental variable models rely on strong assumptions, the alternative causal explanations proposed by Bailey and Bailey and Rind also rely on assumptions—assumptions that appear inconsistent with empirical evidence from data simulations and further examination of the NESARC data.

## Appendix: Details of the Simulations

To investigate the causal structure proposed by Bailey and Bailey, we looked at the case of same-sex identity in men, with stepparent before age 5 as the instrument, as stepparent before age 5 was least likely to be affected by reporting bias. Because most of the statistical mediation found in our data was by childhood sexual abuse, we examined sexual abuse as the mediator. We used existing genetic studies to estimate the likely effect sizes of a given single-nucleotide polymorphism (SNP) on a behavioral outcome. Evidence from genome-wide association studies (GWAS) of anthropometric measures, diseases, and behavioral traits indicate that a given SNP typically accounts for less than 0.5 % of the variation in a trait (Vrieze, Iacono, & McGue, 2012). A recent GWAS meta-analysis suggested that SNPs that affect personality have small or very small effect sizes. This study examined 2.5 million SNPs from more than 17,000 persons and failed to identified even one SNP with GWAS-level significance for neuroticism; effect sizes for SNPs associated with openness and conscientiousness were small and not well replicated (de Moor et al., 2010).

We simulated data from 15,000 individuals (in StataIC 11), using assumptions that would produce the largest confounding by gene while still being somewhat plausible given current understandings of genetics. Although we consider many of the assumptions below unlikely, assumptions that we considered *likely* clearly would not support Bailey and Bailey's hypothesis. Our goal with this simulation was to assess whether even these very extreme assumptions would be consistent with Bailey and Bailey's hypothesis:

- We assumed that mother's neuroticism followed a normal distribution.
- We randomly assigned a neuroticism risk allele to the mother with a minor allele frequency (MAF) of 0.2. We assumed the allele increased neuroticism by 0.48 SDs (the maximum effect size found in the GWAS meta-analysis of all personality traits). We note that this combination of effect size and MAF resulted in 3.8 % of the mother's neuroticism being accounted for by this SNP, 7 times greater than 0.5 % estimated for a typical SNP (Vrieze, Iacono, & McGue, 2012).
- We assumed that mother's neuroticism accounted for 25 % of the probability of her child having a stepparent by age 5 (likely to be an overestimation of this effect). We coded individuals with the highest probability of having a stepparent as having a stepparent so that the prevalence of having a stepparent by age 5 was 2.6 %, as in the NESARC dataset.
- If the mother had the neuroticism risk allele, we assigned the risk allele to the child with a 0.5 probability.
- We assumed the child's risk allele for neuroticism increased his probability of having a same-sex identity by 0.48 SDs (the maximum effect size found in the GWAS meta-analysis of all personality traits). We assigned same-sex orientation to men with the highest probability of having a same-sex orientation such that the prevalence was 1.9 %, as in the NESARC data. In the resulting dataset, the SNP explained 3 % of the child's likelihood of having a same-sex orientation, which would be an exceptionally large effect. In the only large population-representative twin study of sexual orientation, genetic effects in total were estimated to explain .34–.39 of the variance in male sexual orientation (Langstrom, Rahman, Carlstrom, & Lichtenstein, 2010). Thus, the neurot icism SNP would explain 8 % of the genetic component of same-sex orientation. This approach also assumes that the gene has the same effect size on neuroticism and sexual orien tation, which is highly unlikely.

Using data resulting from this simulation, we fit a model for same-sex orientation using stepparent as the predictor. The odds ratio (OR) for stepparent in this model was 1.07 (95 % confidence interval [CI] = 0.5, 2.2). In contrast, in the NESARC data having a stepparent was a strong predictor of sexual orientation (OR = 1.8, 95 % CI = 1.2, 2.7).

Since our initial assumptions did not produce the associations found in the NESARC data, we further explored the assumptions required to produce those associations. We assumed that the SNP had an effect size of 1 (presence of the risk allele increased the mother's neuroticism by 1 SD, which resulted in the gene accounting for 14 % of mother's neuroticism). These assumptions resulted in mother's neuroticism accounting for 38 % of the likelihood of having a stepparent before age 5. It seems very unlikely that neuroticism (or any other genetic factor) could account for more than one-third of the risk of divorce or death of spouse and remarriage by the child's age 5. Nonetheless, these assumptions still did not create an association between same-sex sexuality and having a stepparent as large as that in the NESARC data (OR = 1.4, 95 % CI = 0.7, 2.6). To obtain an association similar to that found in NESARC, we assumed the mother's neuroticism increased likelihood of having a stepparent by 1.35 SD, resulting in her neuroticism accounting for 50 % of the likelihood of having a stepparent, a very implausible scenario.

We then turned to the issue of statistical mediation by childhood sexual abuse. We assumed that the child's underlying risk of sexual abuse (a continuous variable) was a function of mother's neuroticism, such that mother's neuroticism increased risk by 0.3 SD and the child's risk gene increased risk by 0.48 SD (following Bailey and Bailey's hypothesis that the child's gene would affect the child's experience of sexual abuse more strongly than the

mother's neuroticism). With these somewhat arbitrary assumptions, mother's neuroticism accounted for 10 % of the child's risk of sexual abuse and the child's neuroticism risk allele accounted for 5 % of the child's risk of sexual abuse (an exceptionally large, and unlikely, effect size).

We assigned sexual abuse as high, medium, low or none based on risk of abuse to match the prevalence of sexual abuse in NESARC, irrespective of sexual identity. With this assumption, the prevalence of moderate and high levels of sexual abuse in gay men were substantially lower than these prevalences in NESARC and sexual abuse did not mediate the association between stepparent and likelihood of being gay. We therefore next assumed that the child's nascent sexual identity affected risk of abuse. We assigned sexual abuse as high, medium, low or none according to risk of abuse to match the prevalence of abuse among persons with and without same-sex identity in the NESARC data. We then calculated ORs for same-sex identity as the dependent variable with having a stepparent before age 5 and sexual abuse (high, medium, low or none) as the independent variable. In this model, the association of stepparent with same-sex identity was attenuated from the model without sexual abuse (adjusted model, OR = 1.2, 95 % CI = 0.6, 2.2; unadjusted model OR = 1.7, 95 % CI = 0.9, 3.0). These results were similar to those obtained using the NESARC data.

**STATA code**

```
*15000 observations
clear
set obs 15000
*minor allele frequency=0.2
set seed 2829382
*does the mother have the allele?
gen gene=uniform()>.8
*gene increases neuroticism by .48 standardized beta
*(maximum effect from Big 5 genetic study)
gen momneurotic=invnorm(uniform())+(.48*gene)
reg momneurotic gene
*gene accounts for 3.8% of mother's neuroticism
*producing a prevalence of 2.6% of people with stepparent before age 5
*makes mom's neuroticism account for 25% of the probability of having a stepparent
gen stepparent=(0.8*momneurotic+invnorm(uniform()))>2.62
sum
*does the child inherit the allele from the father?
set seed 1462964
gen childhasgene=uniform()>.9
*does the child inherit the allele from the mother?
gen coinflip=uniform()>.5 if gene=1
replace childhasgene=gene if coinflip=1
tab gene childhasgene, r col
*child's orientation: 0.019 of men are gay in NESARC, use maximum effect from Big 5 study
gen childgay=invnorm(uniform())+(0.48*childhasgene)>2.2
*with these assumptions, the neuroticism gene accounts for 3.3% of the child's likelihood of being gay
logit childgay childhasgene
*does this set of assumptions produce an association between child's orientation and stepparent that we see in the data? (no, no association)
sum
tab childgay stepparent, chi2 column exact row
*does it produce an OR=1.8, as we see in the data? (no, OR=1.07)
logit childgay stepparent, or
*what if the gene increases neurotic by 1 SD and childgay by 1 SD instead?
gen momneurotic1=invnorm(uniform())+(1*gene)
*producing a prevalence of 2.6% of people with stepparent before age 5
gen stepparent1=(0.8*momneurotic1+invnorm(uniform()))>2.75
sum
*mother's neuroticism now accounts for 29% of the probability of having a stepparent before age 5
logit stepparent1 momneurotic1
*child's orientation: 0.019 of men are gay in NESARC
*using 1 SD effect of the gene on orientation
gen childgay1=invnorm(uniform())+(1*childhasgene)>2.47
tab childgay1
*the gene now accounts for 14.6% of the probability of the child being gay
logit childgay1 childhasgene
*does this set of assumptions produce an association between child's orientation and
*stepparent that we see in our data? (no, OR=0.9)
tab childgay1 stepparent1, chi2 column exact row
logit childgay1 stepparent1, or
*what if mother's neuroticism accounts for a larger portion of likelihood of having a stepparent?
gen stepparent2=(momneurotic1+invnorm(uniform()))>3.05
sum
*mother's neuroticism now accounts for 36% of the probability of having a stepparent<age 5
logit stepparent2 momneurotic1
*does this produce the association between child sexual identity and stepparent in NESARC?
*No, OR=1.4
tab childgay1 stepparent2, chi2 column exact row
logit childgay1 stepparent2, or
```

*what if mother's neuroticism accounts for an even larger portion of likelihood of stepparent?

gen stepparent3=(1.35*momneurotic1+invnorm(uniform()))>3.75

sum

*mother's neuroticism accounts for 50% of the likelihood of having a stepparent

logit stepparent3 momneurotic1

*does this produce the association between child sexual identity and stepparent in NESARC?

*almost, OR=1.7, 95% CI=0.9, 3.0

tab childgay1 stepparent3, chi2 column exact row

logit childgay1 stepparent3, or

*adding abuse

*abuse risk a function both of mom neuroticism and the child's gene

gen childabuse=invnorm(uniform())+(.3*momneurotic1)+(.48*childhasgene)

*mother's neuroticism accounts for 10% of child's sexual abuse risk

reg childabuse momneurotic1

*child's gene accounts for 4.7% of his risk of sexual abuse

reg childabuse childhasgene

*if childgay does not affect risk of sexual abuse in this simulation, sexual abuse among gay men

*here (low, 2.2%; medium, 3.1%, high, 3.1%) is far lower than in NESARC (low, 2.2%; medium, 4.3%, high, 7.1%)

gen sexabuse=(childabuse>2.35)+(childabuse>2.05)+(childabuse>1.85)

tab sexabuse

tab childgay1 sexabuse, r col

*and sex abuse does not attenuate the association between stepparent and gay as in NESARC

*adjusted OR=1.6, 95% CI=0.9, 2.9

egen byte sexabuse1=anycount(sexabuse), values(1)

egen byte sexabuse2=anycount(sexabuse), values(2)

egen byte sexabuse3=anycount(sexabuse), values(3)

logit childgay1 stepparent3 sexabuse1 sexabuse2 sexabuse3,or

*making sexual orientation affect sexual abuse

*prevalences in NESARC: straight men: low (1.8%), medium (1.7%), high abuse (2.0%)

*gay men: low (1.9%), medium (4.7%), high abuse (12.6%)

drop sexabuse sexabuse1 sexabuse2 sexabuse3

gen sexabuse=(childabuse>2.35)+(childabuse>2.05)+(childabuse>1.85) if childgay1==0

replace sexabuse=(childabuse>1.65)+(childabuse>1.49)+(childabuse>1.34) if childgay1==1

tab childgay1 sexabuse, r col

*does sex abuse attenuate the association between stepparent and gay as in NESARC?

*yes, adjusted OR=1.2, 95% CI=0.6, 2.2

*making indicator variables

egen byte sexabuse1=anycount(sexabuse), values(1)

egen byte sexabuse2=anycount(sexabuse), values(2)

egen byte sexabuse3=anycount(sexabuse), values(3)

logit childgay1 stepparent3 sexabuse1 sexabuse2 sexabuse3, or

## References


Andersen, J. P., & Blosnich, J. (2013). Disparities in adverse childhood experiences among sexual minority and heterosexual adults: Results from a multi-state probability-based sample. *PLoS ONE, 8*, e54691.

Bailey, D. H., & Bailey, J. M. (2013). Poor instruments lead to poor inferences: Comment on Roberts, Glymour, and Koenin (2013). *Archives of Sexual Behavior, 42*, 1649–1652.

de Moor, M. H., Costa, P., Terracciano, A., Krueger, R., De Geus, E., Toshiko, T., et al. (2010). Meta-analysis of genome-wide association studies for personality. *Molecular Psychiatry, 17*, 337–349.

Langstrom, N., Rahman, Q., Carlstrom, E., & Lichtenstein, P. (2010). Genetic and environmental effects on same-sex sexual behavior: A population study of twins in Sweden. *Archives of Sexual Behavior, 39*, 75–80.

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., et al. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature, 460*, 748–752.

Rind, B. (2013). Homosexual orientation—From nature, not abuse: A critique of Roberts, Glymour, and Koenen (2013). *Archives of Sexual Behavior, 42*, 1653–1664.

Roberts, A. L., Glymour, M. M., & Koenen, K. C. (2013). Does maltreatment in childhood affect sexual orientation in adulthood? *Archives of Sexual Behavior, 42*, 161–171.

Saewyc, E. M., Skay, C. L., Pettingell, S. L., Reis, E. A., Bearinger, L., Resnick, M., et al. (2006). Hazards of stigma: The sexual and physical abuse of gay, lesbian, and bisexual adolescents in the United States and Canada. *Child Welfare, 85*, 195–213.

Vrieze, S. I., Iacono, W. G., & McGue, M. (2012). Confluence of genes, environment, development, and behavior in a post Genome-Wide Association Study world. *Development and Psychopathology, 24*, 1195–1214.