

Exploring Model Specification Options for Adolescent Gender Expression

Developed in Women, Gender and Health 207:
Advanced Topics in Women, Gender, and Health
Harvard T.H. Chan School of Public Health, Spring 2018

Course Instructed by Allegra R. Gordon and Avanti Adhia

Teaching Example Authored by Emily Unger and Gabe Murchison

Appropriate Courses:

This teaching example was developed for PHS 2000A, the core quantitative methods course for the Population Health Sciences program at the Harvard T. H. Chan School of Public Health. The example may be useful for other courses that cover regression at an introductory or intermediate level.

Background:

In many regression scenarios, inappropriately specifying the relationship between an exposure and an outcome can lead to incorrect conclusions. Misspecification may result from unfamiliarity with alternative specification options, as well as from unexamined assumptions about the exposure and outcome at hand. In choosing a model specification, it is important to both visualize the data and to think critically about one's theoretical framework and the substantive questions it entails.

Gender is frequently included in regression models in the population health sciences. Though nearly always modeled as binary (male/female), gender is in fact a multidimensional construct; two important dimensions are *gender identity* (internal sense of being a boy/man, girl/woman, or another gender) and *gender expression* (the extent to which a person's grooming, mannerisms, preferences, and other traits match those their culture associates with males and/or females). In many cultures, the sex a person is assigned at birth—typically based on the newborn's anatomy—is expected to dictate both their gender identity and their gender expression. In practice, however, people may have any combination of these characteristics.

In some cases, we might expect masculine or feminine gender expression *per se* to affect health or health behaviors. In other cases, the relevant exposure might be whether the person's gender expression is perceived as appropriate for their current gender identity and/or their sex assigned at birth. People whose gender expression is perceived as less typical for their gender or sex assigned at birth are sometimes described as *gender non-conforming* or (especially among

children and adolescents) *gender-expansive*. Broadly, this includes men and boys perceived as “androgynous” or “feminine” and women and girls perceived as “androgynous” or “masculine.”¹

Being gender-expansive has been identified as a risk factor for mental health problems among sexual minority (e.g., lesbian, gay, bisexual) adolescents, likely due to stigma and harassment (D’Augelli, Grossman, & Starks, 2006). However, few studies have examined how gender expression relates to health in the broader adolescent population. Furthermore, gender expression has typically been modeled as binary (i.e., gender-expansive versus not) or ordinal, potentially obscuring more complex patterns across a spectrum of masculine/feminine gender expression. Overcoming both of these limitations, a 2016 report titled *Health Risk Behaviors among Gender-Expansive Students* (Gill & Frazer, 2016) used data from the Youth Risk Behavior Surveillance System (YRBSS) to identify relationships between an ordinal measure of gender expression and key health indicators among U.S. high school students.

The data in this report is derived from four major municipalities included a question about gender expression when administering the YRBSS in their high schools. The question was phrased: “A person’s appearance, style, dress, or the way they walk or talk may affect how people describe them. How do you think other people at school would describe you?” with response options very feminine, mostly feminine, somewhat feminine, equally feminine and masculine, somewhat masculine, mostly masculine, and very masculine. Additionally, students were asked “What is your sex?” with response options male and female. It is important to note that these data cannot be used to identify transgender students: some transgender students may have responded to the “sex” question by naming their sex assigned at birth, while others may have responded by naming their gender identity. As such, although the data are presented separately for the “male” and “female” groups, one cannot assume that this variable has a one-to-one correspondence with either students’ sex assigned at birth or their gender identity.

This teaching example is based on the report’s findings on the relationship between gender expression, self-reported sex, and the probability of having experienced persistent sadness or hopelessness in the preceding year. Data in the present example were simulated to closely match the report’s results.

Students’ Prior Knowledge:

Students should already be comfortable interpreting coefficients from linear and logistic regression models, including interaction terms. Students should also have been introduced to the notion of modeling linear, categorical, and polynomial relationships in regression, either previously or as part of the same lesson.

¹ Being gender-expansive is not the same as being *transgender*, which describes people whose gender identity (see above) does not align with expectations for their sex assigned at birth. The YRBSS does not routinely include a measure of transgender status. Gender expression is also distinct from sexual orientation. Sexual minority adolescents are more likely than heterosexuals to rate themselves as gender-expansive, but most gender-expansive adolescents are heterosexual (Gill & Frazer, 2016).

Learning Objectives for Students:

After discussing the example, students should be able to:

- Interpret linear, categorical, and polynomial terms in a linear and/or logistic model
- Appropriately select linear, categorical, and/or polynomial specification in light of both the data and the theory or research question at hand
- Understand that multiple dimensions of gender, including gender expression, may interact to affect health outcomes, and that non-binary and even non-linear specification may be required to understand these relationships

Teaching Methods:

This interactive example can be incorporated into a lecture on model specification. Based on the background information and plots presented, students are asked to identify the technical and substantive advantages and disadvantages of each model. The module is designed to take approximately 15 minutes.

References:

D'Augelli, A. R., Grossman, A. H., & Starks, M. T. (2006). Childhood Gender Atypicality, Victimization, and PTSD Among Lesbian, Gay, and Bisexual Youth. *Journal of Interpersonal Violence, 21*(11), 1462-1482. doi:10.1177/0886260506293482

Gill, A., & Frazer, M. (2016). *Health Risk Behaviors among Gender Expansive Students: Making the Case for Including a Measure of Gender Expression in Population-Based Surveys*. Washington, DC: Advocates for Youth. Retrieved from <http://advocatesforyouth.org/storage/advfy/documents/YRBSS.pdf>