

# Potentially faulty data spotted in surveys of drug use and other behaviors among LGBTQ youth

August 18 2023, by Joseph Cimpian

---



Credit: Pixabay/CC0 Public Domain

Federal data on LGBTQ student health [contain a significant amount of potentially exaggerated or untruthful responses](#), raising questions about

how they might skew people's understanding of risky behavior among teens. These inaccuracies affect some responses more than others. That's according to an analysis my colleagues and I did of high school surveys administered by the Centers for Disease Control and Prevention, better known as the CDC.

Without accounting for this invalid data, the CDC results suggest that for every heterosexual boy who uses steroids, three LGBQ boys use steroids. After accounting for the invalid data, neither group is shown to use steroids more. In contrast, disparities for being bullied or considering suicide were not affected by potentially invalid data.

Over 12,800 [high school students](#) during the 2018-2019 school year reported whether they identified as LGBQ—that is, lesbian, gay, bisexual or questioning—or heterosexual on the national [Youth Risk Behavior Survey](#). They also responded to items related to their health and well-being.

We first estimated what the risk disparities between LGBQ and heterosexual youth were before accounting for potentially invalid data. We then used a machine-learning algorithm to detect response patterns that suggested when youth were providing extreme or untruthful responses.

For example, we treated their responses with suspicion if they reported eating carrots four or more times every day and said they were impossibly tall. That means we gave less weight to their responses when we re-estimated all of the disparities. We then saw how the disparities changed after the potentially invalid responses were taken into account.

After accounting for invalid data, disparities in [drug use](#)—including steroids—injecting drugs, cocaine, ecstasy and pain medication without a prescription were not as pronounced. LGBQ boys appeared to use

injected drugs four times as often as heterosexual boys. But after accounting for the likely invalid data, neither group was more likely to use injected drugs.

Yet, while some outcomes were susceptible to invalid data, others were not. For example, LGBTQ boys and girls were about twice as likely to be bullied at school and two to three times as likely to consider suicide. This shows that not all outcomes are equally affected by invalid data.

## Why it matters

The Youth Risk Behavior Survey provides vital information on the health and behaviors of high school students. It informs research regarding [teen sexual behaviors, drug use and suicide risk](#).

[Our study](#) and others using different methods to account for invalid data [consistently find that](#) LGBTQ students are at a much higher risk for being bullied and for suicide, consistent with [CDC reports](#) on these outcomes.

It is critical to address the [ongoing stigmatization that LGBTQ+ people face](#) to reduce these mental health disparities. Yet, when researchers don't check for invalid data, they might conclude that other differences are larger and more deserving of attention and resources than they are.

Policymakers and researchers must ensure that large-scale data collection efforts have safeguards for data quality.

We asked the CDC for a comment on our study's findings. In response, they directed our attention to an [FAQ page](#) that discussed validity and reliability in a general sense. The CDC's response did not specifically address the issue of how invalid data can have a disproportionate effect on minorities, which is a significant concern raised by our research.

## What other research is being done

Other studies have found that invalid data can disproportionately influence [low-incidence outcomes like heroin use](#) and [minority populations](#), including [adoptees](#), [disabled individuals](#), [racial or ethnic minorities](#), [immigrants](#) and [transgender individuals](#).

Moreover, the issue of invalid data is not confined to youth surveys. Studies examining [public health behaviors](#) during the COVID-19 pandemic and surveys on [sexual orientation among adults](#) have also encountered invalid responses, raising further questions about their accuracy.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Potentially faulty data spotted in surveys of drug use and other behaviors among LGBTQ youth (2023, August 18) retrieved 9 May 2024 from <https://medicalxpress.com/news/2023-08-potentially-faulty-surveys-drug-behaviors.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--