

Violations of "The data speaks for itself"

Let us take an example of receiving feedback through different surveys on a national reform in a country's education system that is going to be in force from the next academic year. The reform is to completely change the existing model of education and implement a different approach. The policy envisages English as a mandatory medium of instruction, broad-based, multi-disciplinary, holistic Under Graduate education with flexible curricula, creative combinations of subjects, integration of vocational education, and multiple entries and exit points with appropriate certification.

An Online survey was rolled out on an independent educational platform to its users asking about their opinion on the new educational reform. The results said around 87.7 % of students are feeling positive about the change and welcoming it.

Here the data is biased because the survey is rolled out to only its users about a national reform, neglecting the students who don't use this platform for their academic needs. Also, it is an online survey, which means it is accessible to students who have the internet and mobile phone/computer accessibility. There are many students in the rural areas of the country who doesn't have access to the internet/technology, their opinion on this reform matters a lot and will add weight to the data because, as they have been being taught in their mother tongue which is now changing into English and they will have a different opinion about it. In this scenario, the statement "the data speaks for itself" is highly problematic and very misleading, since the data received is biased. [Example link](#)

In order to make data more balanced/unbiased, the survey should be conducted both online and offline (In Schools/colleges located in both cities and villages). This survey results will be appropriate than only the Online survey which was conducted before.