

International Baccalaureate

Essay

Theory of Knowledge

“Statistics conceal as much as they reveal.” Discuss this claim with reference to two areas of knowledge.

Word Count: 1597

February 16, 2021

The title postulates the problem of an existing equivalent relationship between a statistic through the words "as much as". This implies that these concepts are almost equivalent in presence, which is a questionable statement. The word "statistic" is seen as "quantified observation", as we tend to discretise and quantify our observations into numbers, whilst statistics allow us to interpret these numbers. "Revealing" suggests what is implied from a statistic. However, this "implication" is not immediately produced, as once said, "correlation does not equal causation". The "implication" from a statistic is ambiguous due to different variations of knowledge it can suggest, and these variations can stem from personal knowledge of the interpreter and/or the intended purpose of the statistic creator. "Concealing", on the other hand, is a definition which is the antonym of "revealing", hence we may define "concealing" as "implications" which are rarely seen or considered. Since statistics are a form of quantified information, its ideas mainly stem from the validity of mathematics and its ability to produce knowledge, therefore this essay will focus on the area of knowledge of mathematics. Furthermore, this essay will explore the idea of "quantification" and "discretisation" for the production of statistics within the area of knowledge of human sciences, and thus explore to the extent to which a statistic can reveal the truth using both mathematics and human sciences.

A statistic today shows that in the United States "Life expectancy at birth for the total population has reached an all-time American high level, 77.5 years, up from 49.2 years at the turn of the 20th century" (CRS, 2006). These "revealed" numbers, however, are to be deduced information from. It is required that these numbers are interpreted by its audience. This increased life expectancy is used today to reinforce the idea that human conditions are improving, that is that we "live better" (Windsor, 2011) or that the past's short life span is "shocking" (Satherley, 2019) within the mass media, which is an interpretation commonly used amongst the population. To the majority, this statistic ultimately "reveals" the improvement of the conditions of human life and safety as implied by the common media, therefore it is also the knowledge that is extracted.

However, this is not the only interpretation that is present from the same statistic of the same numbers. Theodore John Kaczynski, a mathematics professor, argues that this statistic does not imply improved “physical security” of the current person in the modern age. In his perspective, these statistics “conceal” information, that is information that could vastly impact these numbers. Such information includes modern disasters that are seen rarely or disasters whose impacts will be seen in the future. These examples include but are not limited to nuclear wars, biological mass murder weapons, pandemics, and global warming. The difference in the “revealed” and thus “interpreted” information between Kaczynski’s viewpoint and that of the common media implies that the revealed numbers of a statistic can result in different knowledge, the culprit being the differing perspectives of the interpreter. This effect is similar to confirmation bias, as interpreters tend to use figures and statistics to support their respective claims from their personal knowledge and thus “concealing”. Hence, it is possible to argue that a statistic is similar to a literary text - flexible for opinion - given that a statistic must be judged with the person’s own reasoning and perspective. This leads to a side effect that all unique interpretations and implications of a statistic are “concealed” knowledge as these are driven by personal knowledge that the vast majority do not have.

The creation of a statistic ultimately relies upon mathematics and its quantification. This concept over time has substantially raised concerns with the idea that statistics and numbers not “revealing” any “true” knowledge, specifically highlighted by interpretivists in Human Sciences. Within sociology, it is argued that human behaviour is simply too difficult to quantify, therefore any produced statistic in an attempt to explain human behaviour oversimplifies the truth. Whether the sociological experiment is done through qualitative interviews or quantitative surveys, we can take an example of the Bobo Doll experiment from an interpretivist viewpoint. The Bobo Doll experiment that was designed to see media can induce violence in children and be done using a qualitative method - namely interpretive observation - to conclude its findings (McLeod, 2014). However, if we consider a “quantified” version of this experiment, would we only count how many times a child punches the doll? What about other forms of possible violence e.g. swearing, grabbing, pushing? How do we

quantify the severity of violence applied? If a statistical and a quantified replica of the experiment was conducted, the number of factors that could classify as “violence” will be limited, therefore simplifying and hence concealing the “truth”. Hence, complex scenarios, which indeed may not be just limited to human behaviour and human sciences in general, may not be the best method to “reveal” the “truth”. In fact, the interpretivist viewpoint in human sciences could arguably stretch to history as past events can only be described with words rather than numbers. For example, a death toll for a specific population within a specific war would answer how many, but would not answer how, when, why, or who, similar to how a Bobo Doll statistic would fail to discretise emotion such as violence.

The problem with statistics also arises with the hidden uncertainties and methodology. For example, during the beginning of the COVID-19 pandemic, the UK has reported an official statistic revealing the total number of cumulative infected cases to be 17,089 on March 28 (Stewart, n.d.). However, the statistics published by Imperial College London contradicts this. The university predicted that the likely number of infections lies between the range of approximately 0.8-3.5 million people the same day (Seth Flaxman, 2020, p.6). Is then possible to argue that the great uncertainty within the methodology of statistics is always up for questioning. The statistic produced by Imperial College London is based on other countries’ observations, whose specific observations may not apply to the COVID-19 cases of the UK. Similarly, the statistics produced by the local government may not be representative of the real number, given that the testing supplies were limited, thus not representative of the whole population. It can be argued that such uncertainties and limitations arise due to the observational nature of statistics, that is, statistics do not take into account what is not observed. Therefore, if this “concealed” knowledge is taken into account to further produce knowledge, it can contradict the current observations due to its flaws (as with the case of the example above). However, it is impossible to escape this “concealment” as it is not possible to take into account everything. Hence, it is impossible for statistics to “reveal” everything by the nature at which they’re produced.

However, simplism and reduction of information is not necessarily a negative trait within the production of knowledge from statistics. The book “Number and Numbers”, by Alain Badiou, argues that the process of one’s label, identity, and surrounding discoveries can be very quickly established and is done so through statistics and mathematics: namely “What is it to exist, if not to assert oneself through a favourable account? In America, one starts by saying how much one earns, an identification that has the merit of honesty” (Badiou, 2008, p.7). Our desire to understand and create knowledge can be done quickly using mathematics, whether it is modeling our surrounding physical phenomena using mathematics e.g. the movement of a particle, or human behaviour e.g. statistics in sociology or psychology, or even both e.g. COVID-19 pandemic. If the process was too complex, one would spend a lot of time identifying the “merit of honesty” or other everyday chores for which we label ourselves, whether it is being a top percentile within a specific subject in school or even age at which we consider “adolescent”, “teen” or “adult”. Therefore, whilst mathematics may “oversimplify” the knowledge, it may be essential to create quick knowledge for use within everyday life and communication, the method at which statistics are created also.

In conclusion, the production of statistics allows us to produce and “reveal” knowledge quickly, helping assert labels and identities merely. Each statistic produced is ultimately interpreted to produce knowledge. This contradicts the concept of mathematics being completely “objective”, as when a person tries to make sense of the obtained numbers through statistics, it must ultimately be subjected to apply the person’s own knowledge and ultimately make sense of the numbers. In this case, it could be argued that a statistic only “reveals” what a person sees, but “conceals” on a larger scale due to all the perspectives and personal knowledge that they do not have. Furthermore, when mathematics and statistics are utilised to explain complex systems e.g. human behaviour, the idea of “quantifying” information can oversimplify the “truth”, as suggested by interpretivists. This can result in concealment of the reality, as not all information is taken into account for quantification to create the statistic, and hence ultimately not all information is also used for the production of

knowledge. Lastly, statistics are ultimately made up of strictly only observation, and not of unobserved factors, as seen in COVID-19 statistics. Taking into account all such possibilities, the amount that a statistic “conceals” outweighs in comparison to how much a statistic “reveals”. Despite the concealment, however, statistics have allowed us to quickly “reveal” information for our own needs even if they may be limited as to how much they “conceal”, which ultimately lead to the production of knowledge and essential labels such as age. For more complex scenarios, however, statistics alone may not be enough to reveal the truth.

References

- Badiou, A. (2008). *Number and Numbers* [Book]. Polity Press.
- CRS. (2006). *Life Expectancy in the United States* [Report]. Retrieved 2020-02-14, from <https://www.everycrsreport.com/reports/RL32792.html>
- McLeod, S. (2014). *Bobo Doll Experiment* [Text]. Retrieved 2021-02-06, from <https://www.simplypsychology.org/bobo-doll.html>
- Satherley, D. (2019). *Our shockingly short lives without modern medicine revealed* [Article]. Retrieved 2020-02-14, from <https://www.newshub.co.nz/home/lifestyle/2019/12/our-shockingly-short-lives-without-modern-medicine-revealed.html>
- Seth Flaxman, A. G. H. J. T. U. T. A. M. H. C. C. W. H. Z. T. B. J. W. E. M. M. I. C. C.-. R. T. A. C. G. C. A. D. S. R. M. A. C. V. N. M. F. L. C. O. S. B., Swapnil Mishra. (2020). *Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries* [Report]. Retrieved 2021-02-06, from <https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-30-COVID19-Report-13.pdf>
- Stewart, C. (n.d.). *Cumulative number of coronavirus (COVID-19) cases in the United Kingdom (UK) since January 2020* [Text]. Retrieved 2021-02-06, from <https://www.statista.com/statistics/1101958/cumulative-coronavirus-cases-in-the-uk/>
- Windsor, A. (2011). *Medicine will not only make us live longer, but live better* [Article]. Retrieved 2020-02-14, from <https://www.theguardian.com/zurichfutures/story/0,,1952688,00.html>