www.bircu-journal.com/index.php/birci
Budapest International Research and Critics Institute-Journal (BIRCI-Journal)
Humanities and Social Sciences
ISSN 2615-3076 Online)
ISSN 2615-1715 (Print)
Budapest Institute

The Difference between the Use of the Word Influence and Correlation in Quantitative Research and Its Implications

M. I. Suhifatullah[1], Mastur Thoyib[2]
[1,2]*Universitas Islam Syekh Yusuf, Indonesia*
misuhifatullah@unis.ac.id, mthoyib@unis.ac.id

Abstract

Logical and rational ways of thinking are the main characteristics of scientific thought, apart from being systematic, measurable, and factual. But in fact, many scientific works such as theses and lecturers' research that use the words influence and correlation both use a correlational approach in their methodology or there is no difference. This is a denial of logical and rational thought, even though it may go unnoticed out of ignorance. This paper is expected to provide a correct understanding of the differences in the use of the word influence and correlation in quantitative research and its implications for research methodology.

Keywords
differences; influence word; correlation word; quantitative research; implications

I. Introduction

Research activities are inseparable from academic culture or scientific thinking in understanding problems and overcoming them based on the power of reason and logic. That is what distinguishes the ordinary way of thinking from the academic way of thinking. Humans have the ability to reason, which means to think logically and analytically. The ability of humans to think logically, because they have a language to communicate their abstract thoughts, so that humans not only have experiences in the form of knowledge, but are also able to develop their experiences and knowledge (Suhartono, 2005). It is necessary to realize that something logical is usually easy to understand by our reasoning but something that is illogical usually conflicts with our minds and hearts (Sobur, 2015).

Reason plays a big role in human life, without human reason tends to be less than optimal in thinking and has difficulty planning something accurately. So, only by using human reason can activities, and work to organize life in a directed and measured manner (Fuadi, 2016). Current human resource development should focus on thinking skills, which involve critical, systematic, logical, and creative thinking. In a process of thinking activities require understanding and the ability to reason about problems related to something that is being thought (Somatanaya, 2017).

However, in school and college learning there is a tendency emphasize more on the form of memorization than the learning process that involves logical thinking. The research Fahinu (2013) findings reveal that the learning process in higher education places a lot of emphasis on learning in the form of memorization, not reasoning, so that students' critical thinking skills do not develop (Heryadi, 2017). There is a difference between the process of memorizing and logical thinking. Memorization only refers to the achievement of mere memory abilities, whereas logical thinking refers more to understanding (being able to explain), application skills, analytical skills, synthesis abilities, and even evaluation abilities to form skills (a process) (Saragih, 2006).

According to Astuti et al (2019) Education is an obligation of every human being that must be pursued to hold responsibilities and try to produce progress in knowledge and

experience for the lives of every individual. Education is one of the efforts to improve the ability of human intelligence, thus he is able to improve the quality of his life (Saleh and Mujahiddin, 2020). Education is expected to be able to answer all the challenges of the times and be able to foster national generations, so that people become reliable and of high quality, with strong characteristics, clear identities and able to deal with current and future problems (Azhar, 2018).

In the lecture process and the process of guidance for writing scientific papers, lecturers should provide logical thinking guidance to students, not just directing students to completion research reports in the form of a thesis. Logical thinking is systematic reasoning in producing correct decisions (Heryadi, 2017). Correlation and causation are terms which are mostly misunderstood and often used interchangeably. Understanding both the statistical terms is very important not only to make conclusions but more importantly, making correct conclusion at the end (Singh, 2018).

However, in reality, many lecturers start the process of guidance for writing scientific papers, by asking students to submit thesis titles, without asking for an explanation of the core of the problem in his research . In this case there is no logical thinking process, only an instant thought process without a correct understanding of the proposed thesis title, so that many thesis titles are irrelevant and logical in the use of their methodology. In this regard, it is assumed that there are still many lecturers who do not fully understand the research methodology, for example in understanding the differences in the use of the word influence and relationship in quantitative research and their implications for research methodology.

On the basis of the above problems, the purpose of writing this paper is to analyze the differences in the use of the word influence and correlation in quantitative research and their implications for the research methodology.

II. Research Method

This research includes the study of research literature to find references relevant theory to the case or the problems found. References theory obtained by the research literature studies serve as the basic foundation and a major tool for research practice in the middle of the field.

III. Results and Discussion

There are many student research titles or lecturers' scientific works that use the word influence in quantitative research, whereas generally they use correlational data analysis methods. If the data analysis method is correlational, it is more appropriate to use the word correlation. The scientific work of a lecturer thesis and research that uses the word influence or correlation is generally not different or both are correlational. This fact is found in almost every faculty/university that has been uploaded to Google. Ataqila (2016) explained that, in theory, the word correlation serves to measure the degree of closeness (correlation) between two variables based on literature reviews that are related. However, the analysis of the direction of the correlation is not based on a standard theoretical basis. Meanwhile, the word influence serves to explain the causal pattern of a variable that affects (free) to the affected (bound) variable.

Actually, every scientific research title always reflects the problem, as well as the research method that will be used. When the researcher determined the title, "The correlation between variable X and variable Y", the research method used was different

from the research entitled "The effect of variable X on variable Y". Titles that use the word correlation are intended to measure the degree of closeness (correlation) between two or more variables. While the title that uses the word influence is intended to analyze the cause and effect that occurs in an independent variable on the dependent variable. Sulistya (2020) explained that correlation is a pattern of correlation between two things or variables that appear at the same time. There is a fact that every rainy season there are always many accidents. This is an example of a correlation between two events or conditions. But this fact does not show a causal correlation, because the rainy season is not the cause of traffic accidents. Traffic accidents during the rainy season may be caused by slippery roads, many holes in the road are not visible due to waterlogging.

When one variable has a correlation with other variables, it does not automatically mean that it has a causal correlation. Tufte (2006) dan Aldrich (1995) explained that, Correlation does not mean to indicate causation, an expression used in science and statistics to mean the existence of a correlation between two variables, does not automatically mean that one variable is the cause of the other variable, though correlation is an indicator of the presence or absence of a causal correlation. Wikipedia in Mustafa (2011) explains, the phrase's that correlation proves causation is a logical fallacy in that two events occurring together are claimed to have a causal correlation. Madhavan (2019) explained that, Correlation and causality can appear similar. However, recognizing the difference between the two can avoid doing something worthless and of low value and create something people will continue to talk about.

There has been a logical error, when there is a correlation between two variables, it is always interpreted as evidence of a causal relationship. It is true that in a causality there is always a correlation, but in a correlation there is not always a causality. Holland (1986) explained, causation requires not only correlation, but also dependence on comparative facts. Suppose a student is performing poorly on a test and suspects that it is because he or she is not studying, then to prove this a comparative fact is required, the same student wrote the same test in the same circumstances but had studied the night before. In this way, the cause can be observed by comparing version 1 with version 2).

The wrong and illogical way of thinking when stating that there is an influence of a variable on other variables is only based on the Regression statistical formula. True, this formula is used to calculate the magnitude of the influence and also the level of the correlation. However, the influence of one variable on another is not determined by the regression formula, but by the accuracy of the research method. Madhavan (2019) explained, causation and correlation can exist at the same time, but correlation does not imply causation. Causation explicitly applies to cases where action A causes outcome B. On the other hand, correlation is simply a correlation. Event A is related to event B - but one event doesn't always cause another event to occur. Denis and Legerski (2006) explain, there are various kinds of statistical applications designed by statisticians, but the use of statistical analysis in research must have a strong basis in relation to the research design. Because statistics and theory are two things that have different positions and functions. Statistical methods such as path analysis, regression formulas and all statistical tools, are only calculating machines applied to substantive problems of theoretical concern. Although the path analysis model allows one to assess the various indirect effects that might be hypothesized in a correct theory, this does not mean that theory and statistics should be put together or "bridged" in any way.

3.1 Correlational Research

Many students and lecturers in formulating research titles, do not seem to question whether the condition of the dependent variable is caused or influenced by the existence of the independent variable. Maybe they don't even consider the existence of independent and dependent variables important in the title of their research. They just want to know whether there is a correlation between one variable and another. For example, "is there a correlation between learning motivation with student learning achievement?" Another example is "is there a correlation between the level of intelligence with the quality of life?"

Or "the correlation between height with body weight".

If the research objective is to determine whether or not there is a positive correlation or the size of the correlation between variables, then the type of research is correlation. Associated with these objectives is usually done by obtaining data on variables X and Y (if two variables), then the results are processed by correlation analysis. For example, what you want to know is the relationship between Learning Motivation with Student Learning Achievement, so respondents who are the population or sample are asked to fill out a questionnaire about Learning Motivation and look for data about the average value of learning outcomes for each student who becomes the respondent. Then the score for each variable is calculated with a certain statistical formula to determine the level of the relationship. An example of the score for each variable is as follows.

Respondent Number	Learning Motivation (X)	Learning achievement (Y)
1	**98**	**90**
2	**87**	**85**
3	**90**	**85**
4	**77**	**80**
5	**86**	**76**
etc	etc	etc

After calculating, the correlation score is obtained between -1 to 1. If the result is minus it means that the correlation is negative (an increase in variable X, followed by a decrease in variable Y or vice versa). If the result is plus, it means that the correlation is positive (an increase in variable X is followed by an increase in variable Y or vice versa). If the result = 0 or about 0 (0.00 ...) then there is no correlation between variables X and Y. Then based on certain criteria, the size of the correlation can be determined and with certain calculation methods can also be obtained the importance (significance) of the correlation.

When the researcher finds that there is a relationship between the X and Y variables, the researcher can conclude the size of the correlation and whether the correlation is significant or not, but cannot conclude that the value of one variable affects the value of the other variable. In this case, the researcher cannot conclude or interpret that Learning Motivation is a causal variable of the Student Achievement value (causality relationship), because in essence what causes high or low Student Learning Achievement is the student's ability to solve test or exam questions and assignments provided by the teacher, not Learning Motivation, even if it makes sense. Because Learning Motivation is not the cause of Student Learning Achievement. This means that there is no causal (causal) correlation between Learning Motivation and Student Achievement, although there is a positive correlation between the two variables.

In Statistics Knowledge Portal said, Correlation tests for a correlation between two variables. However, seeing two variables moving together does not necessarily mean we know whether one variable causes the other to occur. This is why we commonly say "correlation does not imply causation". Jim Frost (2019) mengatakan, If correlation does not prove causation, what statistical test do you use to assess causality? That's a trick question because no statistical analysis can make that determination. In this post, learn about why you want to determine causation and how to do that.

In other cases, for example, there is evidence that a person's height has a positive correlation with body weight. The taller a person increases the weight. However, these findings cannot say that a person's height is a causative factor to his weight. Because what can cause weight gain is the diet or the type of food consumed.

What needs to be considered is related to the correlation research above, that the correlation of learning motivation as variable X with student learning achievement as variable Y cannot be the opposite. Likewise, the correlation between height as variable X with body weight as variable Y. This is because logically learning motivation as variable X is a contributor to changes in student learning achievement as variable Y, likewise height as variable X is a contributor to change. which occurs in body weight as variable Y, as in the following figure.

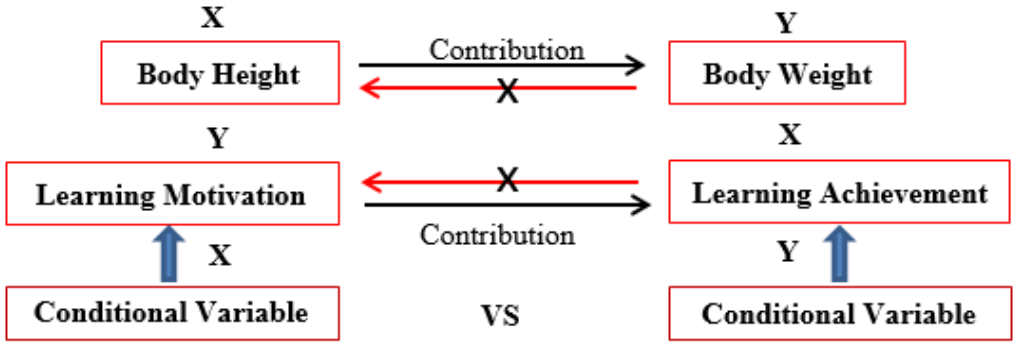

Figure 1. *Characteristics of Correlational Patterns Testing the relationship between conditional with conditional variables*

The picture above shows that correlational research has characteristics where the variable X and variable Y are both conditional variables or outcome variables. However, variable X is usually predicted to be stronger or more dominant in condition so that it can contribute to changes that occur in the condition of variable Y. In the case above, it is certainly not relevant if student learning achievement is positioned as variable X and learning motivation as variable Y, because that becomes the core problem in schools is usually student learning achievement, so that it becomes the Y variable or the dependent variable, although in principle, in correlational research there can be a two-way or reciprocal correlation. However, sometimes the reciprocal correlation is difficult or even unreasonable, such as in the case of the correlation between body height as variable X with body weight as variable Y, of course in this case it cannot be the opposite.

3.2 Causality Research

Research sometimes does not just want to know the correlation between two or more variables, but wants to know the effect of a variable on other variables or to find out which variables are the cause of changes in the condition of other variables or a causal correlation. To find out this, the types of research or scientific efforts carried out include conducting experiments or a series of controlled experiments. For example, what you want

to know is the effect of fertilizer application on banana tree growth, then it is necessary to carry out various experiments in the laboratory or in the "field". Because the growth of banana trees is not only influenced by the action of applying fertilizers, it is necessary to control any other variables which are thought to have an effect on the growth of banana trees, so that the results of the effects are convincing or unbiased. Campbell and Stanley in Williams (2007) support experimental research designs with a high level of control over various other variables that can influence, so that the validity of the results can be accounted for, and this is the real experiment. While the quasi-experimental research design involves nonrandom research objects, this can result in very limited control to of other variables, so that experiments like this are unlikely to produce high validity.

In various scientific studies to find out the effect, impact or influence of a certain treatment on changes that occur in the object of treatment - for example the effect of drug use on healing disease - it must be an experimental study that is carried out repeatedly. Statements or testimonies of a person or group of people related to the influence or effectiveness of drugs, for example about the efficacy of green coconut water in curing certain diseases, in the medical world are unacceptable.

In the https://research-methodology.net/causal-research/ account which was downloaded on 14/08/2021, it was revealed that the core of causality research is

Causality research has a commitment to explain certain problems or situations that have a causal relationship between variables. In causality research, the most popular method for collecting primary data is through experimentation. Because the claim of a causal relationship can only be accepted, if specifically there is strong causal evidence. It is important to know, there are three elements as causal evidence, namely:

a. In causality, the element that becomes the cause must first exist or occur before there is an effect. Cannot claim, that the increase in sales of a product as a result of improved promotions, if the increase has occurred before the improvement of promotions.
b. There must be a logical and rational relationship pattern between the two variables. Cannot claim that employee performance increases due to the influence of organizational culture, if management never changes its organizational culture.
c. It is ensured that the causal relationship between the two variables must be true and not because there are other variables. There should be no 'third' variable associated with the two variables being studied. This means that in an experiment there must be control over other variables that can affect the dependent variable.

Experimental research has long been a scientific tradition, especially in the exact sciences. Because in this research, manipulation of independent variables can be carried out which can change or bring up the dependent variable. This is important, because sometimes research does not just want to know whether there is a correlation between an independent variable and the dependent variable. But also to find out with certainty the existence of a variable that causes the occurrence or emergence of the dependent/dependent variable. For example, what causes employees to be stressed, what causes students' learning motivation to decrease, what causes consumers to be dissatisfied, Is it true that the decline in student learning motivation is caused by inappropriate learning methods? Is it true that employee motivation is low due to low wages or salaries? and so forth. What he wants to know is what variable causes all the negative things to appear or occur, with the aim of being able to provide a solution to the problems mentioned above.

In the context above, the type of research that should be used must be different from research whose purpose is "merely" to find out the correlation between two variables. If the main objective is to find out the existence of cause-effect (cause-effect) in a certain setting, then the type of research is an experimental design or a causal-comparative design.

However, it is necessary to be careful with non-experimental causal-comparative research. As revealed by Sukadji (1997: 33), in many non-experimental studies, direct control of independent variables is not possible, and manipulation or random assignment is also impossible. The hypothesis which states "if X then Y, when there is no control over the possibility of other X variables, then the findings of a hypothetical correlation between X and Y cannot be confirmed with confidence as in experimental situations. So non-experimental research can basically be said to have inherent weakness, namely the absence of control over the independent variables.

If a researcher wants to know how much or whether there is an effect of the independent variable on the dependent variable, then inevitably the researcher must apply an experimental or "causal-comparative" research design by controlling the other independent variables tightly. But in fact many researchers are wrong, for example for the purpose of research entitled "The effect of smoking on cancer", researchers usually take a number of samples. Then the sample (patients suffering from cancer) was measured on the frequency of smoking (x-independent variable) and also measured the severity of the cancer (y-dependent variable). After that the value of x is associated with the value of y, using the regression formula. The result is then interpreted that the variable x has/has no effect on y. Or it can be concluded that patients suffering from cancer are caused/not caused because they smoke, or smoking is the cause/not the cause of cancer

The conclusion from the research steps as above, clearly contains a logical error "cum hoc ergo propter hoc", because the research method used is correlational, namely by connecting the x data (smoking frequency) and y data (cancer), then with the regression formula the data x and data y are calculated to find out how big the level of the relationship is. Connecting the values of two different variables, of course, is not a process to find out whether or not there is an influence, impact, effect of smoking on cancer

A logical step to find out whether smoking causes cancer with the research title: "The effect of smoking on cancer" then the research method must be experimental, not correlational. The steps taken were to find a group of volunteers (respondents) who really did not suffer from cancer (research sample), then they were quarantined so that they did not eat other foods that could cause cancer (control variable). In a certain period of time they are required to smoke. After that, each sample or respondent was examined for cancer. If the researcher wants to use a regression formula, then the x data is the cancer of each respondent before smoking and the y data is the cancer of each respondent after smoking. So the data x and y are the same variable data (cancer) from the same person (respondent). Pretest – Treatment – Posttest. The calculation results will show how much influence smoking has on cancer. It can be very small to very large, or from insignificant to significant.

So, when wanting to know whether there is a relationship between smoking habits (variable X) and cancer (variable Y), the researchers must have data related to smoking habits variables - for example measured by the number of cigarettes smoked - and data related to the severity of cancer from respondents who have a smoking habit. Then the two data are connected by using certain statistical formulas. When the results show a significant relationship (the more cigarettes smoked, the higher the severity of the cancer), but it still can't be said that the severity of cancer is caused by smoking habits, although it may make sense. Therefore the appropriate research title is "The Relationship Between Smoking Habits With Cancer", not "The Effect of Smoking on Cancer"

IV. Conclusion

The use of the word influence in quantitative research should be for research that connects process/treatment variables with conditional variables or the results obtained, for example: the effect of the incessant promotion of certain companies (process/treatment) on buyer interest (results or conditional), the effect of giving discounts on certain products (process/treatment) on the sales value of certain products (results or conditional), the effect of Welfare Improvement (process/treatment) on employee performance (results or conditional), etc. However, it must be remembered that the methods and data analysis used are not correlational (connecting data x with data Y), but must compare the data before the process or treatment with the data after the process or treatment.

The use of the word correlation in quantitative research should be used for research that relates conditional variables to conditional variables, for example: the correlation between customer satisfaction (conditional) and customer loyalty (conditional), or the correlation between organizational culture (conditional variable) and employee performance (conditional variable).

So, there is no prohibition against giving the title "The Effect of X on Y", but it is necessary to consider the consequences of the title in the research method. Only experimental research is feasible to apply if it intends to find out whether the value of one variable (X) causes or has an effect/impact/effective on changes in the value of another variable (Y). If it turns out that the research method is not experimental but correlational, the title of the research should be changed to "The Relationship between X and Y". Even then, if the nature of the variable is conditional with conditional, not process against conditional.

References

Aldrich, John (1995). "Correlations Genuine and Spurious in Pearson and Yule" (PDF). Statistical Science. **10** (4): 364–376. doi:10.1214/ss/1177009870

Artaqila. 2016. https://artaqila.blogspot.com/2016/09/perbedaan-pengaruh-dan-hubungan-dalam-penelitian.html?m=1

Astuti, R.W., Waluyo, H.J., and Rohmadi, M. (2019). Character Education Values in Animation Movie of Nussa and Rarra. *Budapest International Research and Critics Institute-Journal (BIRCI-Journal)*. P. 215-219.

Azhar,A. (2018). Students' Trends in Islamic Communication Postgraduate in2010-2016 State Islamic University of North Sumatera (UINSU). *Budapest International Research and Critics Institute (BIRCI-Journal)*, P.206-214.

Daniel J. Denis and Joanna Legerski, (2006). Causal Modeling and the Origins of Path Analysis. Theory & Science (2006) ISSN: 1527-5558, https://theoryandscience. icaap.org/ content/vol7.1/denis.html

Fuadi. 2016. Fungsi Nalar menurut Muhammad Arkoun. *Substantia, Volume 18 Nomor 1, April 2016.*

Heryadi, Dedi. 2017. Menumbuhkan Karakter Akademik Dalam Perkuliahan Berbasis Logika. *Jurnal Pendidikan Karakter, Tahun VII, Nomor1, April 2017*

Holland, Paul W. (1986). "Statistics and Causal Inference". *Journal of the American Statistical Association*. 81 (396): 945–960. *doi*:*10.1080/01621459.1986.10478354*

https://research-methodology.net/causal-research/

Jim Frost (2019). Causation in Statistics: Hill's Criteria. https://statisticsbyjim.com/basics/causation/#comments

Madhavan, Archana. 2019, Correlation vs Causation: Understand the Difference for Your Product, https://blog.amplitude.com/causation-correlation

Mustafa, Hasan. (2011). Hati-hati dengan judul penelitian yang menggunakan kata "Pengaruh, Efek, Dampak X Terhadap Y", http://hasanmustafa.blogspot. com/2011/08/ penelitian-kausal-dan-korelasi.html

Prima Sulistya 2020. https://mojok.co/prm/komen/versus/beda-korelasi-dan-kausalitas-agar-tak-salah-ambil-kesimpulan/

Saleh, A., Mujahiddin. (2020). Challenges and Opportunities for Community Empowerment Practices in Indonesia during the Covid-19 Pandemic through Strengthening the Role of Higher Education. *Budapest International Research and Critics Institute-Journal (BIRCI-Journal)*. Volume 3, No 2, Page: 1105-1113

Saragih, Sahat. 2006. Menumbuhkembangkan Berpikir Logis dan Sikap Positif terhadap Matematika Melalui Pendekatan Matematika Realistik. *Jurnal pendidikan dan kebudayaan Departemen Pendidikan Nasional. Badan Penelitian dan Pengembangan, Edisi Juli 2006*

Singh, Seema. 2018. Why correlation does not imply causation? https://towardsdatascience. com/why-correlation-does-not-imply-causation-5b99790df07e

Sobur, H. A Kadir. 2015. Logika dan Penalaran dalam Perspektif Ilmu Pengetahuan, *TAJDID Vol. XIV, No. 2, Juli-Desember 2015*

Somatanaya, A.A. Gde. 2017. Nalisis Kemampuan Berfikir Nalar Matematis Serta Kontribusinya Terhadap Prestasi Belajar Mahasiswa. *Jurnal Teori dan Riset Matematika (TEOREMA) Vol. 1 No. 2, Hal, 55, Maret2017ISSN 2541-0660 © 2017*

Statistics Knowledge Portal, https://www.jmp.com/en_au/statistics-knowledge-portal/what-is-correlation/correlation-vs-causation.html

Suhartono, Suparlan. 2005. *Sejarah Pemikiran Filsafat Modern*. Yogyakarta: Ar Ruzz Media.

Sukadji, Soetarlinah. 1997. Awas: Studi Kausal-Komparatif. *Buletin Psikologi, Tahun V, Nomor 2, Desember 1997, ISSN : 0454 - 7108*

Surat, I Made. 2016. Pembentukan Karakter Dan Kemampuan Berpikir Logis Siswa Melalui Pembelajaran Matematika Berbasis Saintifik. *Jurnal EMASAINS Volume V, Nomor 1, Maret Tahun 2016 ISSN 2302-2124*

Tufte, Edward R. (2006). The Cognitive Style of PowerPoint: Pitching Out Corrupts Within *(2nd ed.).* Cheshire, Connecticut*:* Graphics Press. ISBN 978-0-9613921-5-4

Wikipedia, https://www.wikiwand.com/en/Correlation_does_not_imply_causation

Williams, Carrie. 2007. Research Methods. *Journal of Business & Economic Research – March 2007 Volume 5, Number 3*