

# Fulfilling the law of a single independent variable and improving the result of mathematical educational research

H Pardimin<sup>1</sup> and N Arcana<sup>1</sup>

<sup>1</sup>Pendidikan Matematika, Universitas Sarjanawiyata Tamansiswa

E-mail: nyoman.arcana@ustjogja.ac.id

**Abstract.** Many types of research in the field of mathematics education apply the Quasi-Experimental method and statistical analysis use t-test. Quasi-experiment has a weakness that is difficult to fulfil "the law of a single independent variable". T-test also has a weakness that is a generalization of the conclusions obtained is less powerful. This research aimed to find ways to reduce the weaknesses of the Quasi-experimental method and improved the generalization of the research results. The method applied in the research was a non-interactive qualitative method, and the type was concept analysis. Concepts analysed are the concept of statistics, research methods of education, and research reports. The result represented a way to overcome the weaknesses of quasi-Experiments and T-test. In addition, the way was to apply a combination of Factorial Design and Balanced Design, which the authors refer to as Factorial-Balanced Design. The advantages of this design are: (1) almost fulfilling "the law of single independent variable" so no need to test the similarity of the academic ability, (2) the sample size of the experimental group and the control group became larger and equal; so it becomes robust to deal with violations of the assumptions of the ANOVA test.

## 1. Introduction

The success of teachers in teaching is influenced by many factors. One factor is the method of teaching. Therefore, many students of the Faculty of Education who are writing their thesis conduct researches on a comparison of two or more methods of teaching. In the field of Mathematics education, for example, students do a comparison between the effectiveness of Example and Non-Example method with the conventional method in learning geometry. The research method used is usually a Quasi-experimental method. Some examples of researchers using quasi-experimental research methods are studies conducted by [1,2]; this research aims to know the difference of effectiveness between two teaching methods in learning Mathematics. The quasi-experimental method is used because it is not possible to do sampling using simple random sampling directly on students to form experimental group and control group. This is because students have been formed in the classroom (intact group).

The quasi-experimental method is a research method that uses all students who have formed in one intact group as a sample, an example is a group of students in one class [3]. Therefore, the samples are taken in the class, not the students [4]. The quasi-experimental method according to its definition, is weak in its randomization. The word "quasi" in Latin means almost or near, so a quasi-experiment can be interpreted as "near-experiment" [5].

This method requires the researcher to identify the control group as similar or as close as possible to the experimental group in the underlying characteristics [6], such as academic ability and student motivation. Since it is difficult to find two "similar" classes, the main problem faced by the quasi-experimental method is that it is difficult to meet a "law of the single independent variable". The law of



a single independent variable is a law essentially stating that all conditions in the experimental class must be the same as the control class except the teaching method (Mill, 1986 in [7]).

The data in the research [1,2] were analyzed using a t-test of two mean differences. The significance of the mean differences between the two sets of data can be tested using the t-test of two mean differences [8]. However, the t-test applied to investigate the significance of the difference in effectiveness between the two teaching methods has weaknesses. The disadvantage is that the conclusions obtained through the t-test are incapable of seeing whether differences of effectiveness occur in groups of students with high, moderate or low academic ability [9]. So the conclusions of the research obtained through the t-test are less extensive in its generalization.

From the above description, it appears that the quasi-experimental method and t-test has a weakness. The weakness of the use of quasi-experimental research method is difficult to fulfill "the law of a single independent variable," a law stating that the only one difference is allowed to exist between the two classes, that is the teaching method. The weakness of using the t-test is only able to see if there are differences in the effectiveness of the classical but is unable to see how the differences in each group of students' academic ability (the ability of high, medium and low). In fact, the difference in academic ability is very necessary to be considered in the learning of Mathematics.

This research aims to find ways to reduce the weaknesses of the quasi-experimental method and improves the generalization of the research results.

## 2. Method

The research applied non-iterative qualitative method, also called analysis research. The type of analysis used was the concept analysis. The concepts analyzed were the concept of statistics, educational research methodology, and Mathematics education research report.

The data were collected with a literary technique. This technique collects data by reviewing documents and archives on matters related to the research. The data sources were the documents.

The data analysis was carried out by identifying, analyzing, and synthesizing concepts from the collected data. Then, the interpretation of the results of the synthesis was carried out to discover a new concept.

## 3. Results and Discussion

The results of this research are the concept of the relationship between quasi-experimental method and the law of single independent variable; the concept of the relationship between quasi-experiment and Balanced Design; the concept of relationship among t-test, Factorial ANOVA, and the generalization of research result; concept of the robust of the F-test, concept of Balanced-Factorial Design, and the procedure of applying Balanced-Factorial Design.

There are two general categories of experimental research methods, namely: true experimental and quasi-experimental [5]. Quasi-experimental research may be regarded as the best effort of experimental research when it is impossible or unreasonable to meet all the criteria of a true experimental. In the true experimental method, the study subjects (students) are taken randomly, while on quasi-experiment randomization is done to the class.

As a result of randomization performed in the class, there are many problems in interpreting the observed difference between groups [10]. In Mathematics for example, perhaps most of the experimental class backgrounds are better at Math than the control class. It shows that quasi-experiments are weak in the fulfillment of the fundamental similarity of characteristics between the experimental class and control class.

Basically, the experimental method is based on two assumptions about variables, [7] one of which is: if the two situations are the same in all respects except for the variables are added or removed from one situation, then the differences later emerged between the two situations can be associated with these variables. This statement is called "the law of a single independent variable".

The core of the experimental method is control. One of the variables that must be controlled is the variable of students' academic ability. The academic ability of the students in the class whose learning using  $X_1$  method should be the same as the students' academic ability in the class that got  $X_2$  method.

One way to achieve this similarity is by rotating classes in intervals during the experiment. Suppose the experimental period is divided into two equal intervals of time, the materials will be taught are also divided into two units, namely unit 1 and unit 2. The complexity of the material in unit 1 has to be balanced with that of unit 2. During the first interval material unit 1 is taught, one class (class A) applies  $X_1$  method and the second class (class B) applies  $X_2$  method. During the second interval unit, 2 material is taught, but class A applies method  $X_2$  and class B applies  $X_1$  method. If such rotation is done then it is said that research applies counterbalanced or Balanced Design [7]. Table 1 is a Balanced Design that is used to compare the two methods of teaching, namely  $X_1$  and  $X_2$ .

**Table 1.** The Balanced Design with Two Teaching Methods

Period interval & the subject matter	$X_1$ teaching method	$X_2$ teaching method
Period interval I, Unit 1	Class A	Class B
Period interval II, Unit 2	Class B	Class A
Combination	Score $X_{1A} +$	Score $X_{2B} + X_{2A}$
Mean	$X_{1B}$ Mean of score $X_1$	Mean of score $X_2$

Based on Table 1,  $X_{1A}$  score is combined with  $X_{1B}$  score. Likewise, the  $X_{2A}$  score is combined with the  $X_{2B}$  score.

Testing assumptions of normality and homogeneity of variance are required by the t-test and the ANOVA (Analysis of variance) test since both of these tests include the parametric test. There are 4 assumptions for parametric tests [9,11], those are: (1) normally distributed data, (2) the variance between homogeneous groups, (3) the data interval, and (4) the independent data. If these assumptions are not met, then, the conclusions of ANOVA will deviate from what it should be. However, fortunately, in general, the consequences caused by not fulfilling assumptions with respect to ANOVA is not too bad [12]. In other words, a moderate deviation from the terms outlined in assumptions is not too dangerous. For example, if there is a slight deviation from normal assumption or assumption of homogeneity of variance it would only affect very small on the test results and the conclusions.

ANOVA (also called F test) includes a robust test, that is, it will be not too problematic if the assumption is not met [9]. If the sample size is large, then the sampling distribution will have a normal distribution (based on Central Limit Theorem). If the sample size is the same, F test is very robust against violations of normality; thereby also, F test remains accurate even though there is a slight violation of the homogeneity assumption. However, if the sample size is not the same and it results in a violation of the homogeneity assumption, the conclusions derived from the F test can be biased.

The above statement is supported also by [11], for samples of the same size or almost the same, ANOVA is less sensitive to non-fulfillment of normality and homogeneity of variance. This is a beneficial use of Balanced Design. Even many modern statistical experts feel that the testing of assumption is only necessary for small samples [13].

One-Factor ANOVA and t-test have proved to be very useful in some studies of the exact sciences, such as Agriculture and Pharmaceuticals, but poor when they are used in the field of science concerning behavior, for example in the field of Mathematics teaching [14]. Because One-Factor ANOVA will not be able to reveal the interactive effect between teaching methods and student academic ability.

The analysis which can reveal this interaction is Factorial ANOVA. The term "way" or "factor" is often used to describe the number of independent variables measured by an ANOVA test. The design is called a factorial design. This design was first developed by R.A. Fisher in Agriculture [14]. Another

advantage of using a factorial design is that it increases the significance of the differential effectiveness of two methods of teaching [8]. Besides, it can also increase the generalizability of research results.

Supposedly, there is a question on whether the effectiveness of a method of teaching also depends on other factors, such as students' academic ability factor. Such question can be answered by using ANOVA Factorial. The effectiveness of a teaching method can be seen from the students' learning achievement (achievement of test scores). If the learning achievement is seen from two factors, for example, factors of teaching methods and level of academic ability, then Two Factors ANOVA is applied to analyze the data. If the teaching method consists of 2 (two) levels ( $X_1$  and  $X_2$  method), while the academic ability of students consists of 3 (three) levels (High, Medium, and Low), then the design is called a 2 x 3 Factorial Design. Table 2 below is 2 x 3 Factorial Design.

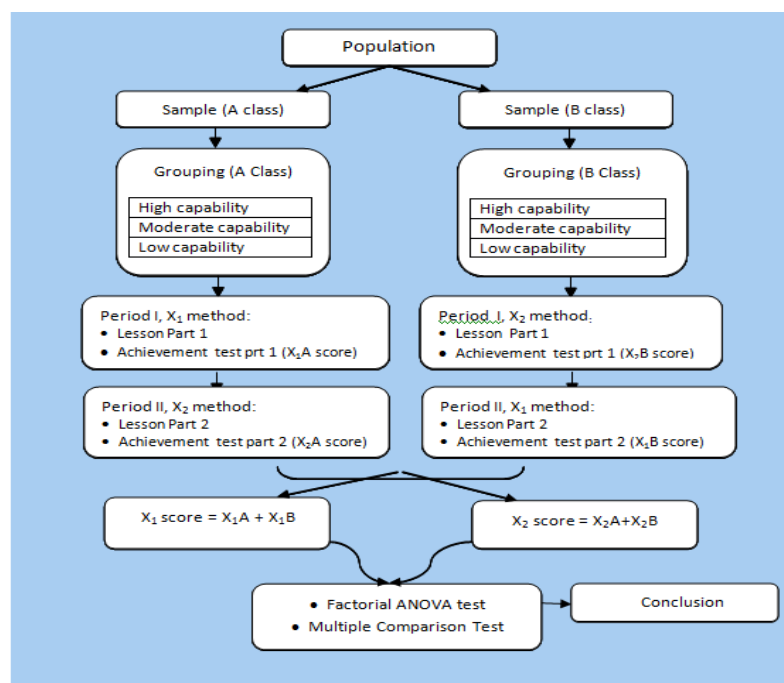
**Table 2.** The 2 x 3 Factorial Design

Academic Ability	$X_1$ Teaching method	$X_2$ Teaching method	
High	AT score <sup>a)</sup>	AT score	Average row
Moderate	AT score	AT score	Average row
Low	AT score	AT score	Average row
	Average column	Average column	

*Note* <sup>a)</sup> AT score is Achievement Test score

Based on Table 2, it appears that researchers can compare test scores for each level of academic ability.

Based on the advantages of Balanced Design and Factorial Design, the authors propose to apply a Factorial-Balanced Design to counteract the weaknesses of the quasi-experimental method and the t-test. Factorial-Balanced Design is a combination of Factorial Design with Balanced Design. This design name is the authors' own creation. The procedure for carrying out the research on the use of the Factorial-Balanced Design is presented in Figure 1. This procedure has been practiced [15], in fact, the application of this procedure is not simple, but the results provide information that is unbiased and it strengthens generalization.



**Figure 1.** Cart of procedure research

Figure 1 shows that two classes were taken as samples. The sampling method used cluster random sampling, and therefore it obtained the class A and class B. Class A was divided into three ability groups, those were groups of high, moderate and low. The basic grouping was based on a summative test score and it has been obtained before the research was done. Grouping was also performed in class B.

The subject matter was divided into two units, namely unit 1 and unit 2. The experimentation time was divided into two, namely the first period and the second period. The first period was taught unit 1 material: class A got X1 method, while class B got X2 method.

The first period ended with achievement test for unit 1. The second period was taught unit 2 material: class A applied X2 method, while class B applied X1 method. The second period ended with achievement test for unit 2. Achievement test scores as a result of applying the X1 method in class A were merged with those in class B. As well, achievement test scores as a result of applying the X2 method in class A were merged with those in class B. The form of a merger is shown in Table 2.

The data in Table 2 were analyzed using Factorial ANOVA test followed by multiple comparison tests. The multiple comparison tests used was the Scheffe test. Scheffe test is one of the multiple tests most frequently used [11]. It is even guaranteed as the best test [8]. This test has several advantages, one of which is that a pair can be selected in accordance with the desired pair.

#### Conclusion

The law of single independent variable is difficult to be fulfilled by the quasi-experiment method. The approach of the fulfillment of this law can apply Balanced Design. Through this design, the class is rotated in intervals during the experiment so that each student is paired with him/herself. The advantages of this design are: first, the sample size in both groups becomes larger and equal; this condition increases the robustness of ANOVA against the assumptions of normality and homogeneity of variance; second, it does not need to test the similarity of average academic ability between groups because each student is paired with him/herself.

The use of t-test cannot conclude whether the effectiveness of the teaching method is the same at every level of academic ability of students, so the generalization is less extensive. A Factorial Design is applied to determine whether the effectiveness of teaching methods depends on the level of academic ability. Based on the previous explanation, a way to fulfill the law of a single independent variable and increase generalization of mathematics education research is the application of a Factorial-Balanced Design. The design is a combination of Factorial Design and Balanced Design. The procedure of applying the Factorial-Balanced Design is not simple, but the results provide information that is unbiased and it strengthens generalization.

#### 4. Conclusions

The law of single independent variable is difficult to be fulfilled by the quasi-experiment method. The approach of the fulfillment of this law can apply Balanced Design. Through this design, the class is rotated in intervals during the experiment so that each student is paired with him/herself. The advantages of this design are first, the sample size in both groups becomes larger and equal; this condition increases the robustness of ANOVA against the assumptions of normality and homogeneity of variance; second, it does not need to test the similarity of average academic ability between groups because each student is paired with him/herself.

The use of t-test cannot conclude whether the effectiveness of the teaching method is the same at every level of academic ability of students, so the generalization is less extensive. A Factorial Design is applied to determine whether the effectiveness of teaching methods depends on the level of academic ability. Based on the previous explanation, a way to fulfill the law of a single independent variable and increase generalization of mathematics education research is the application of a Factorial-Balanced Design. The design is a combination of Factorial Design and Balanced Design. The procedure of applying the Factorial-Balanced Design is not simple, but the results provide information that is unbiased and it strengthens generalization.



## References

- [1] Rochimawati N and Kusmanto B 2016 *Union* **4** 1 pp 79-88
- [2] Alfiah Y and Arigiyati T A 2015 *Union* **3** 2 pp 169-175
- [3] Aznam N, Sumarno and Rahmat A 2005 *Metodologi Penelitian untuk Peningkatan Kualitas Pembelajaran, Penelitian Kuasi Eksperimen dan PPK* .Jakarta: Depdiknas pp 4-5
- [4] McGowan H M 2011 *Journal of Statistics Education* **19** 2 pp 4-5
- [5] Bradley K 2009 *Quasi Experimental Research, A Comparison to Experimental Research* .The Gale Group, Inc. All rights reserved
- [6] White H and Sabarwal S 2014 *Quasi-experimental Design and Methods, Methodological Briefs: Impact Evaluation* **8** Florence UNICEF Office of Research 2
- [7] Ary D, Jacobs L C, Sorensen C and Razavieh A 2010 *Introduction to Research in Education.*, 8<sup>th</sup> ed USA: Wadsworth, Cengage Learning pp 267-319
- [8] Moortgat L R 1980. *Analysis of Variance from t-Test to Analysis of Covariance* (Manila: De la Salle University) pp 28
- [9] Field A 2009 *Discovering Statistics Using SPSS, 3th ed* (London: SAGE Publications Ltd) p 133
- [10] Fife-Schaw C 2006 *Research Methods in Psychology* Quasi- experimental Designs-Semantic Scholar<https://pdfs.semanticscholar.org> pp 79-80
- [11] Deauna C M 1982 *Applied Statistics for Education Psychology and the Social Sciences, 1<sup>st</sup> ed* (Manila: De la Salle University)
- [12] Sudjana 1985 *Desain dan Analisis Eksperimen* (Bandung: Tarsito) pp 48-49
- [13] Hayes W L 1973 *Statistics for the Social Sciences, 2<sup>nd</sup>* (New York: Holt, Rinehart & Winston)
- [14] Best J W 1981 *Research in Education* (USA: Prentice-Hall, Inc) pp 58
- [15] Ngubaedilah A, Arcana I N and Suyadi A 2016 *Efektivitas Model Example Non Example dengan Game Puzzle Terhadap Prestasi Belajar Matematika Pokok Bahasan Geometri pada Siswa Kelas VII SMP Negeri 2 Mlati* (Yogyakarta: Universitas Sarjanawiyata Tamansiswa)

## Acknowledgments

The authors would like to extend their gratitude to Ahmad Ngubaedilah who has been practicing Balanced-Factorial Design (design created by the author) in his research for his thesis in Mathematics Education of Sarjanawiyata Tamansiswa University, Yogyakarta, also to Th. Laksmi Widyarini, M.Hum., for having proofread this article.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.