EVALUATION OF THE ABILITY GROUPING METHOD ON THE EFFECTIVENESS OF TEACHING AND LEARNING IN MTs STUDENTS. X CIANJUR

Silmina Fauziyah[1], Lintang Septiana Yohan[2] & Sabila[3]
e-mail : silminafauziyah@gmail.com[1], lintangseptia93@gmail.com[2], sabilacloud@gmail.com[3]
Faculty of Psychology
Program Study Masters of Psychology Profession
Persada Indonesia University Y.A.I. Street Jakarta **INDONESIA**

ABSTRACT

The aim of this study was to find out the effectiveness of ability grouping method on teaching and learning process of MTs. X students. The subjects were 251 students (program and regular class students VII, VIII, IX) and 53 teachers. The research design used the evaluation research with quantitative and qualitative approach (mix method). The measuring tool used CIPP model evaluation scale (24 items) and effectiveness scale (41 items), as well as interviews conducted to the headmaster and some students as supporting data. Descriptive data processing obtains that, the evaluation of good teacher evaluation is 58,5% and student 62,5%. Input evaluation, 68% teachers and 59% students. Process evaluation, 70% teachers and 52% students. Product evaluation teacher 62% and students 54%. The result of spearman rank correlation test found that the ability grouping method is effective used in teaching and learning process for students (H_0 is rejected, Zhitung = 4,265 >Ztabel = 1,960).

Keywords: Evaluation, Effectiveness, Ability Grouping.

INTRODUCTION

Law No. 20 of 2003 concerning the National Education System, Article 1 states that: "Education is a conscious and planned effort to create a learning atmosphere and learning process so that students actively develop their potential to have religious spiritual strength, self-control, personality, intelligence, noble character. as well as the skills needed by himself, society, nation and state". (Prayitno, 2009, p. 259). The goal of education is actually an effort to develop students both in terms of knowledge, abilities and student character (Yero, 2002). To achieve this goal, each school or other educational institution certainly has different methods.

One method is to group students to study based on their abilities or what is called the ability grouping method. Those who agree with this method argue that a selection process will enable teachers to direct students with different abilities more effectively and allow those who are more capable to achieve higher standards (Ireson & Hallam, 2003). Ireson and Hallam (2003) also added that the Ability grouping method is one solution to the problem of low student achievement.

One school that applies this method is MTs. X which is located in Cianjur Regency. MTs. X is a private junior high school which is part of a foundation which consists of several levels of education ranging from Raudhatul Athfal to Madrasah Aliyah. However, RA, MI, MTs and MA have different developments. After the MTs principal changed in 2009, the principal realized at the beginning of his term of office the quality of the MTs. X both infrastructure and other things are not good and there are very few interested students. The achievements he

obtained were not very many, in fact very few. Most of the students are prospective students who did not graduate from state junior high schools in the area. Finally, the school made various changes and improvements to improve school quality and achieve effective learning. This effort was made to support the school's vision, namely, "to create people who are strong in IMTAQ, strong in science and technology, intelligent in thinking, polite in attitude and wise in acting". One of the changes made is to implement a program and regular class system since 2012, or a program called Ability grouping.

Ability grouping is interpreted from many different points of view by education experts. However, these definitions generally complement each other and refer to one thing in common, namely how to group students who have the same or similar needs, interests or abilities into the same group (Clark & Zimmerman, 1994). Kulik (1992, as quoted in Clark & Zimmerman, 1994) defines the ability grouping method as "separating students who have the same level into one different group or class which indicates differences in ability in school" (h.xii).

Each school has differences in the technical implementation of this method. In fact, the purpose of this ability grouping method is to provide opportunities and convenience for teachers or instructors to be able to deliver teaching material directly to students according to the students' cognitive abilities (Seel, 2012). This is also in accordance with the initial aim of carrying out this method in MTs. X. According to the teachers, this is an effort to increase achievement and improve the quality of their students, as well as a reward for students who excel. Seel (2012) also said that educators believe that if they provide a curriculum that is too easy for students, students will be more easily bored and undisciplined. However, if you provide a curriculum that is too difficult, it will ultimately make students become discouraged and want to end the learning process prematurely. With this ability grouping method, it is hoped that these two negative effects can be avoided, namely by formulating a curriculum and teaching methods that are appropriate to students' cognitive abilities (Seel, 2012).

There are several types of ability grouping methods. Several main types of ability grouping methods used in schools, namely (a) streaming (tracking), (b) Appeal, (c) setting (regrouping), (d) mixed ability (heterogeneous grouping), (e) within class ability grouping , (f) cross-age grouping (cross-grade grouping), (Ireson & Hallam, 2003). MTs. X in this case uses a streaming type ability grouping (tracking) method. Where this type is a grouping of students based on test results which are the students' general abilities or intelligence and is the most rigid type among the other types, because it is based on the assumption that individuals have a fixed intelligence score, which will predict their achievements in all aspects, and this can be measured with objective tests (Ireson & Hallam, 2003).

This student grouping system then has positive and negative impacts on the school. According to the student affairs section, the positive impacts felt by the school include that school performance, especially in outdoor competition activities, has increased quite rapidly over the last five years. Based on data from the student affairs department, the increase in school achievement through competitions held from sub-district to provincial levels in both academic and non-academic fields was approximately 50% after implementing the ability grouping method.

Even when compared to academic competitions, non-academic competitions such as arts and sports are still the ones with the most achievements. In addition, because students with poor moral records are placed in the same class category, namely the regular class, they can easily be found if new cases occur and the school can easily provide treatment, so that cases of

delinquency that occur both inside and outside schools decreased quite a bit, although not significantly.

Ireson and Hallam (2003) also realized that there were negative effects in implementing this method, especially with this type of streaming. According to him, students may become labeled or stigmatized by teachers or other students, which then allows for dissatisfaction or hostility at school (Ireson & Hallam, 2003). MTs student affairs section. Dr.

Students, especially students in regular classes, are less motivated to achieve more because students who are truly cognitively intelligent gather in program classes. According to several students, the teacher's attitude seemed more pleasant in the program class, he was less angry and the material provided was more challenging. Learning materials are also completed more quickly because of their good grasping power, as well as more frequent exams given in program classes.

Apart from that, if the school will face any competition, especially competitions in the academic field such as the Olympics, the school will immediately take representatives from the program class without selection. However, according to the student affairs section, because their abilities are above average, students in program classes tend to look more individual than students in regular classes. The student affairs department added that around 5% of the students in the program class insisted on being transferred to regular classes because they felt they could not stand the competition in the program class. Another reason is because you want to get a ranking and that can be obtained in regular classes.

After five years of running this program, the school has never carried out a specific program evaluation or improvements in the following year. Evaluation is carried out based on teacher reports and activities as well as discussions at each work meeting, the content of which only discusses the work evaluation of teachers, staff and heads of affairs. Not many student complaints are specifically received by the school, except from a few teachers who are quite close to the students. Some students in the program class even forced the student affairs department to move them to regular classes because they felt they couldn't stand the competition in the program class and were bored with having too few friends, especially male friends.

Evaluation is very important to do. Evaluation is defined in the Oxford Advanced Learner's Dictionary of Current English (AS Hornby, 1986, as quoted in Arikunto & Safruddin, 2014) evaluation is "to find out decide the amount or value" (h.1). What this means is that evaluation is an effort to determine a quantity and value which in the process requires caution, is carried out responsibly and can be accounted for and is a process using strategy (Arikunto & Safruddin, 2014). Arikunto (2015) added the definition of evaluation as "an activity to collect information about how something works, which is then used to determine the right alternative in making a decision" (h.2).

Stufflebeam (2017) states that evaluation is a process of describing, obtaining, providing, and applying descriptive and judgmental information about the benefits and feasibility of an objective, implementation design, and results of an object to lead to improvements, provide accountability reports, inform institutional/institutional decisions. dissemination, and improve understanding of the involvement of a phenomenon. The school said that the ability grouping method they implemented was the best method that could be implemented at this time with all

the positive impacts that had been felt over the last five years. However, the comprehensive evaluation process that should have been in place has not been carried out.

The evaluation used in this research is using the CIPP evaluation model from Daniel Stufflebeam. CIPP is an abbreviation of the four dimensions of evaluation according to Kellaghan & Stufflebeam (2003), namely context evaluation, input evaluation, process evaluation and product evaluation.

Context evaluation assesses needs, problems, and opportunities in defining the environment. This evaluation looks at the preparation and readiness of MTs. X In the application of the Ability Grouping method in which there is an evaluation of the background of the implementation of the abilities of grouping, specific objectives and general objectives of the implementation of the abilities of grouping, the basis of consideration of the system applied in the ability of abilities grouping, to the influence of the method of the school budget.

Input evaluation is used to assess competitive strategies and approaches chosen in work plans and budgets in program implementation. This evaluation looks at the extent of the availability of input that will support the running of the ability grouping method and as a solution to problems regarding obstacles and needs in this method. Things that are used as evaluation material include the availability of teaching staff (teachers) and the quality of both whether they support the implementation of this method or not. The quality of students accepted, especially students in program classes and then students in regular classes. how the facilities and infrastructure are ready to support the implementation of the method and the readiness of the school in other matters such as learning materials, learning time, and where the budget used comes from and how it is used.

Process evaluation is used to monitor, document and assess activities. This evaluation looks at the process that occurs in the ability grouping method applied by MTs. X, including the suitability of the process of implementing the method with the system that has been made, the ability of educators (teachers) in the application of this method and the possibility of the method to be continued or not with the conditions of the teaching and educators, the use of facilities and infrastructure in the process whether it is maximal or not yet , as well as obstacles encountered during the process of implementing the method.

Product evaluation is used to identify and assess short-term or long-term results, whether desired or undesirable. This evaluation looks at the achievement of the specific goals and general objectives that have been previously determined, what student needs have been met by implementing this method, the impact obtained by students and teaching staff (teachers) with this method, the obstacles felt during the implementation of the method and the achievements or what undesirable goals occur in applying this method.

Bramley (1996) states that one of the purposes of evaluation is to provide feedback on the effectiveness of the method being used. Komariyah & Triatna (2004, as quoted in Supardi, 2015) stated that effectiveness is "a measure that states the extent to which targets/objectives (quantity, quality and time) have been achieved" (h.2). Abin (1999, as quoted in Supardi, 2015) also added that effectiveness is the correspondence between the results achieved in the form of an achievement or output that can be observed and the expected results in the form of objectives, targets, intended outputs in accordance with what has been determined.

Slavin (1994, as quoted in Supardi, 2015) divides the main elements of effective teaching into four parts, namely QAIT (Quality, Appropriateness, Incentive, Time). The four parts are quality of instruction, appropriate level of instructions, incentives, and fourth is time. So this research uses these four elements to become a standard for the effectiveness of the teaching and learning process at MTs. X uses the ability grouping method to group its student classes. The four aspects of QAIT effectiveness from Slavin are explained in the following details:

a. Quality of teaching (Quality of Instruction)

   The quality of teaching is the result of the quality of the curriculum and lessons which are the teacher's efforts to convey goals and skills that make students understand more easily (Sukardi, 2015). Therefore, Supardi (2015) added that quality teaching is teaching that is easy for students to understand. This is achieved through the teacher's systematic delivery of the material, clear language, explanations accompanied by related examples, emphasis on essential material, and the use of tools in explaining a concept to link the subject matter with knowledge and learning experiences. that students previously had. The objectives of teaching must also be specific and clear.

b. Appropriate Level of Instructions: The level at which the teacher ensures that students are ready to accept new lessons. So that new lessons can be easily conveyed to students, students must have proficiency or entering behavior so that what is conveyed during teaching is in accordance with the students' abilities (Supardi, 2015). If students' abilities are weak while teaching standards are high, then students will not be able to participate in learning well so they may fall behind in their lessons. Meanwhile, if teaching standards are low, students with high abilities will feel disadvantaged because their abilities are not facilitated. One way to overcome this is to group them according to students' abilities or mastery of one skill or teaching (Slavin, 1994, as quoted in Supardi, 2015).

   Another way is with individual teaching methods, so that students can learn according to their abilities with the guidance of the teacher. However, this will still be ineffective because other students will not be facilitated. Another way to improve learning is with corporative learning, where weak students in a group will be guided by their colleagues in that group (Supardi, 2015).

c. Incentives

   This stage is the stage where the teacher ensures that students have the motivation to complete assignments and learn in the subjects given. Here too is the time for teachers to provide this motivation by providing learning that can attract student interest and be fun for students. So teachers need to use various teaching methods so that students do not feel bored and bored during learning (Supardi, 2015).

   The next way is to provide rewards to students in the form of prizes or praise for students who can master a skill or lesson and giving punishment to students who cannot master a skill or lesson. Another form of reward is by reporting students' progress and behavior to parents every week, or by giving rewards to students who show good progress (Slavin, 1994: 147, as quoted in Supardi, 2015). This method is expected to encourage and motivate students to study and do their homework.

d. Time

   This stage is the stage where students are given sufficient time for the desired subjects. There are two time factors that can influence learning. The first is time allocation, which is the time given by the school to teachers to teach a subject. This time is difficult to change because it has become a decision by the school. The second is engaged time or time-on-task, which is the time teachers use to teach and the time students spend studying to gain knowledge and skills.

This research was conducted for several purposes including, (1) to find out and provide feedback regarding the suitability of the ability grouping method with the initial purpose for which the method was created; (2) to find out and provide feedback regarding input (students, teachers, facilities and infrastructure) in supporting the ability grouping method; (3) to find out and provide feedback regarding the process of implementing the ability grouping method; (4) to find out and provide feedback regarding student learning outcomes (achievement and moral development) in accordance with the initial objectives of establishing the ability grouping method; (5) to determine the effectiveness of the ability grouping method in the teaching and learning process at MTs. X.

RESEARCH METHODS

The research design used in this research is descriptive evaluation research using mixed research methods. Evaluation design is "a framework for the process of carrying out an evaluation and a plan for capturing and utilizing data so that information can be obtained with sufficient precision or hypotheses can be tested appropriately and the objectives of the evaluation can be achieved." (Wirawan, 2011, h.147). Mixed research methods are a combination of quantitative and qualitative methods simultaneously in one evaluation process (Wirawan, 2011). Quantitative and qualitative approaches (mixed method) are used because researchers want to obtain valid, reliable and objective data and information that can complement and enrich research results (Sugiyono, 2014).

The evaluation model used in this research is the CIPP evaluation model from Stufflebeam. This evaluation model basically provides an assessment of four aspects of evaluation, namely context, input, process and product. This model was chosen because it directs the target object of evaluation to the process and input to the results so that the results obtained will be more comprehensive.

There are two variables that will be studied in this research. The first variable is evaluation and the second variable is teaching and learning effectiveness. The research subjects in this study were school principals as stakeholders, students and MTs educators (teachers). X uses the ability grouping method to group students.

Number of MTs student population. X for the 2017-2018 academic year there were 873 students, 53 teaching staff and a school principal. Based on the student population, the number of student samples used refers to the Sugiyono (2014) table, namely from a population of 873 students a sample of 251 students (n= 251) was taken with a degree of error of 5%. The sampling technique used is simple random sampling, where all students in the population have the same opportunity to be selected as the sample (Silalahi, 2012). As for the teacher population, because the number is below 100, the entire population is used as the sample.

DATA COLLECTION TECHNIQUE

This research uses two data collection techniques, namely interviews and questionnaires. An interview is a meeting between two people to exchange information and ideas through questions and answers, so that meaning can be constructed on a particular topic (Sugiyono, 2014). Interviews in this research were used to gather initial information regarding the evaluation of the ability grouping method in order to identify the CIPP evaluation model. The second instrument is a questionnaire which is a data collection method if the researcher already knows what will be measured from the variable of interest and is an efficient data collection mechanism (Silalahi, 2012).

This research used four instruments for measurement, namely the principal's interview guide, student FGD guide (guadline) and two instruments in the form of a Likert scale, namely the CIPP model evaluation scale and the teaching and learning effectiveness scale. Both interview guides, FGD guides, and scales for evaluation, all use the CIPP evaluation model from Daniel Stufflebeam based on four evaluation components, namely: (1) context evaluation, (2) input evaluation, (3) process evaluation and (4) product evaluation.

Principal interview guide consists of 8 questions for context evaluation, 4 questions for input evaluation, 6 questions for process evaluation, 7 questions for product evaluation. Meanwhile, the student FGD guide consists of 3 questions for context evaluation, 7 questions for input evaluation, 4 questions for process evaluation, 8 questions for product evaluation. The Likert scale for the teacher evaluation instrument has 22 items and 24 items for students.

The teaching and learning effectiveness scale is prepared based on Slavin's four QAIT indicators, namely (1) quality of instruction, (2) appropriate level of instructions, (3) incentives, and (4) time (time). The number of items for the teacher instrument is 45 items and 41 items for students.

There are four answer choices provided, namely very suitable (SS), suitable (S), not suitable (TS), and very unsuitable (STS). The items of the first instrument, the evaluation scale, are arranged into one category, namely favorable (+) items, while the items of the second instrument, namely the effectiveness scale, are arranged into two categories, namely favorable (+) and unfavorable (-).

Instrument testing is carried out to find out good instruments that can be used in data collection. Item analysis uses the item-total correlation coefficient (rix). From the evaluation scale of 32 items, 22 items were used for the instrument for teachers and 24 items for the instrument for students. Meanwhile, for the effectiveness scale of 56 items, 45 items were used for the scale for teachers and 41 items for the scale for students.

Validity testing is a test carried out to find out whether a measuring instrument really measures what it wants to measure according to its objectives (Azwar, 2015). The way to determine the validity of the measuring instrument is to use the Spearman correlation technique because the data is in the form of ordinal data.

The result is a validity test for the evaluation scale for teachers in the first aspect, namely context evaluation is 0.881 (very high), input evaluation is 0.791 (high), process evaluation is 0.886 (very high) and product evaluation is 0.881 (very high). while the validity results for students for the first aspect of context evaluation were 0.724 (high), input evaluation was 0.727 (high), process evaluation was 0.775 (high), and product evaluation was 0.876 (very high).

The validity results for the effectiveness scale for teachers on the first aspect of teaching quality are 0.938 (very high), appropriateness of teaching level is 0.846 (very high), incentives are 0.920 (very high) and time is 0.790 (high). The validity of the scale for students is that the first aspect of teaching quality is 0.934 (very high), appropriateness of teaching level is 0.915 (very high), incentives are 0.887 (very high) and time is 0.433 (medium).

The final instrument test is by testing the scale reliability. Tests were carried out using Cronbach's alpha reliability coefficient. Based on calculations, the reliability of the evaluation scale for teachers is 0.727 (high), the evaluation scale for students is 0.710 (high). The

reliability of the effectiveness scale for teachers is 0.738 (high) and the effectiveness scale for students is 0.727 (high).

This research was conducted to determine the effectiveness of the ability grouping method in the teaching and learning process for MTs students. X Cianjur. Therefore, statistical testing is needed to test the effectiveness of this method. The hypothesis used is as follows:

$H_0 : r_s = 0$ The ability grouping method is not effectively used in the teaching and learning process for students.

$H_1 : r_s \neq 0$ The ability grouping method is effectively used in the teaching and learning process for students.

Hypothesis testing is carried out using the Rank Order correlation technique developed by Charles Spearman. This technique is used to find the correlation coefficient between ordinal data and other ordinal data, or interval data that has previously been converted into ordinal data (Bungin, 2014). Test Statistics N > 30 using Z transformation. Value Z_{count} then compared with the value Z_{tabel} with test criteria, namely H_0 rejected if $Z_{count} \geq z_{tabel}$ or $\rho \leq \propto$.

This research uses a quantitative and qualitative approach, so that in the analysis, apart from hypothesis testing using the Spearman rank correlation technique, qualitative data analysis is also carried out. Qualitative data collection carried out was an interview using the FGD (focus group discussion) method. This qualitative data is used as additional data to explain things that cannot be explained in the questionnaire. So that data collection is only carried out on subjects who get the most striking scores (for example the lowest) on research results, especially research results in the form of descriptive data.

The analysis carried out was only a simple analysis that went through three stages. The first is to select and discard data that is not important. The second stage is to group the data according to the aspects and indicators of the variables that will be deepened. The third stage is to draw conclusions from data that has been grouped according to existing aspects and indicators.

RESEARCH RESULTS AND DISCUSSION

The results of this research are used to find out whether the ability grouping method is in accordance with the initial objectives of the method, how input (students, teachers, facilities and infrastructure) supports the ability grouping method, what is the process of implementing the ability grouping method, what are the student learning outcomes (achievement and moral development) in accordance with the initial objectives of establishing the ability grouping method and finally whether the ability grouping method is effective in the teaching and learning process at MTs. X.

The first test is to carry out descriptive analysis using median norms. Analysis is carried out on every aspect of the evaluation. The first aspect is context evaluation. The median score obtained for teachers is 21 and the median score for students is 12. So the norm is that if the score obtained is > 21 for teachers or > 12 for students, then the assessment of the context evaluation is good. However, if the score obtained is <21 for teachers or <12 for students, then the assessment of the context evaluation is not good.

Based on these norms, the assessment results obtained by teachers who said they were good were 58.5% and those who said they were not good were 41.5%. As for the assessments obtained by students, 62.5% of students gave good assessments and 37.5% of students gave poor ratings.

The second aspect is input evaluation. shows that the median score for teachers is 10 and the median score for students is 20. So the norm is that if the score obtained is > 10 for teachers or > 20 for students, then the assessment of the input evaluation is good. However, if the score obtained is <10 for teachers or <20 for students, then the assessment of the input evaluation is not good. Based on these norms, a good assessment from teachers was 68% and an unfavorable assessment was 32%. The assessment from students who gave good grades was 59% and 41% were not good.

The third aspect is process evaluation. The median score for teachers is 10 and the median score for students is 15. So the norm is that if the score obtained is > 10 for teachers or > 15 for students, then the assessment of the evaluation process is good. However, if the score obtained is <10 for teachers or <15 for students, then the assessment of the evaluation process is not good. Based on the norms, a good rating from teachers was 70% and 30% was not good. Meanwhile, the assessment from students who said it was good was 52% and 48% was not good.

The final aspect is product evaluation. The median score for teachers is 20 and the median score for students is 21. So the norm is that if the score obtained is > 20 for teachers or > 21 for students, then the assessment of the product evaluation is good. However, if the score obtained is <20 for teachers or <21 for students, then the assessment of the product evaluation is not good. The result is a good assessment from teachers of 62%, not good of 38%. Meanwhile, 54% of students scored good and 46% did not.

These results show that the majority of respondents gave a good assessment of the ability grouping method applied because the good percentage was greater than the bad percentage in all aspects of both students and teachers. Although the comparison between the percentages of good and not good, especially among students, is not significantly different.

Then the testing continued with hypothesis testing using the Spearman rank correlation test statistic. Based on the test results with the help of the system, it was obtained ρ_{value} of 0.000 and α of 0.05. Then Z_{count} amounting to 4,265 and Z_{table} amounted to 1,960. Based on these results, it is good to calculate using Z_{count} and Z_{table} or calculations using ρ_{value} the result is constant H_0 rejected. This means that the ability grouping method is effectively used in the teaching and learning process for students.

This research uses additional data or supporting data in the form of qualitative data through interviews with school principals and FGD (Focus Group Discussion) activities conducted with 16 students with the lowest evaluation scores (not good). 9 program class students and 7 regular class students. Program class students consist of students in grades 8 and 9, while regular class students consist of students in grades 7 and 9.

The FGD process was carried out in two groups, one group consisting of program class students, after completion of the FGD program class student group continued with the second group, namely the regular class student group. The FGD material is aspects contained in the CIPP evaluation model from Stufflebeam to explore the reasons why students give low (not good) assessments.

Principal Interview Results
Context evaluation. In this aspect, the questions asked are how long the school takes to study the application of the method, who ultimately decides whether or not to apply this method,

what are the basic considerations behind determining the ability grouping method, as well as the general and specific objectives of implementing the method. Then consider the long-term opportunities of implementing this method, the school's achievements before this method is implemented, how many classes can be served with this method, and the effect of implementing the method on the budget.

The first question, the principal said that the application of this method had undergone quite a long review. The decision to apply this method also involved several parties, not only the school principal but also the teacher council at teacher and staff performance evaluation meetings. One of the basic considerations for implementing this classification method is to make it easier to select students when there are competition events and as a differentiator from other schools. The general aim of implementing this method is to raise the name of madrasas in general or spread religion. Meanwhile, the specific aim is to produce superior students. However, to produce superior students, adequate support from facilities and infrastructure is needed.

Before implementing this method, schools did not have adequate facilities and infrastructure, especially regarding the number of students. The budget also has a big influence. Because this school is a private school, of course the budget usage will be different from state schools which receive full assistance from the government. Private schools actually use budgets according to the size of their needs. The main target of implementing this method is actually to produce more outstanding students and raise the name of madrasas in general who have different abilities and qualifications from other schools, because the name of madrasas is often underestimated. So that way the school tries to fulfill and facilitate the community's expectations of the school while still paying attention to Islamic law.

Input evaluation. The questions asked in this aspect are about the quality of teachers who are prepared to implement the method, the quality of staff, the quality of students, the quality of the provision of facilities and infrastructure, the readiness of teaching materials, and changes in the budget after implementing this method. The principal realized that in fact, there was no difference between program and regular classes. The treatment given is the same. The only difference is that students in the program class are academically superior. So it would be better if the program class was simply called superior class.

The school principal stated that there was no difference in teaching staff, teaching materials, facilities and infrastructure, budget (in this case tuition fees for students) for both program and regular classes. The principal also realized that teachers who taught in regular classes should be more qualified than teachers who taught in program classes. Teachers should actually feel proud when their teaching goals are achieved well in regular classes rather than program classes.

The facilities and infrastructure provided continue to be improved, although according to him, to support the class method applied is not adequate because program and regular classes receive completely the same treatment. School budgets do not change significantly because of this method. This happens because there is no difference between program classes and regular classes. So the funds spent are only for operational needs as usual.

Process evaluation. This aspect questions matters relating to the program implementation process, the ability of teaching staff to implement the program, utilization of facilities and infrastructure, as well as obstacles encountered during program implementation. Because the

treatment given to both program class and regular class students is basically the same, there is no difference in the implementation process. It's just that, many teachers are more motivated if they teach in program or regular classes.

The principal also admitted that many teachers felt great after successfully teaching in program classes, even though it should have been the opposite. Teachers don't have to prepare too many things if they teach in program classes. However, for teachers with subjects that rely more on psychomotor skills, students in regular classes are superior.

The infrastructure and facilities provided between program classes and regular classes are the same. So that learning media and other facilities are completely the same and used as intended. Although there are still some things that don't exist and are incomplete, such as a computer laboratory. The absence of these facilities is also one of the obstacles to implementing the method.

The lack of differences in facilities and treatment between program classes and regular classes is also an obstacle. Because the name of the program class is just a name because there is no special program given. Only the students are the students with the highest rankings and grades. So in the process, according to the principal the program class is more suitable to be called a superior class.

Product evaluation. Some of the questions asked in this aspect include the achievement of short-term and long-term goals that have been set, student needs that have been met, the impact that students, teachers and staff have had with the ability grouping method, obstacles felt during the method, and the objectives. or undesirable achievements that occur due to the application of this method.

If you look at the purpose of this classification method, one of them is to make it easier for this competition event to occur. The school experienced quite a rapid increase in terms of both academic and non-academic achievements. In fact, the KKM score is quite stable and has now reached 79 in all subjects for both program and regular classes. This KKM score is the average score obtained from all students in both program and regular classes. This means that, even in regular classes, their scores are starting to be good enough so that the KKM score can rise to 79.

The school principal also said that one of the efforts to improve the quality of students is to provide students with guardians at school who function like academic supervisors at universities. The homeroom teacher will guide and develop students throughout the three years of the student's study period at MTs, no longer the homeroom teacher who can change every new school year. So that each guardian will understand the development of their students well.

Student FGD Results
Context evaluation. For students, all the indicators in this aspect are correct and quite good, it's just that the reason for selecting students in competition events is felt to be inappropriate because both program and regular class students have their own advantages.

Input evaluation. Both program and regular class students agreed that there was no difference in terms of the teacher's teaching performance and the quality of the students, only that in the regular class there were more practice questions. Apart from that, the only learning media that does not exist at all is computers.

Process evaluation. The facilities and infrastructure provided are sufficient except for the prayer room (there is none at all). Another thing is that both program and regular class students have the same learning motivation. It's just that regular class motivation decreases in science and mathematics lessons. For program class students, being placed in the program class is not a reward for their achievements.

Product evaluation. School achievements and student interest continue to increase. It's just that this is not because of the class system. Because it turns out that the championships won by the school mostly come from representatives of organizations which consist of program class students and regular class students.

Based on the results of the FGD, students think that this class system is actually not problematic for students. It's just that by being separated like this, differences emerge between students in program classes and regular classes. Their relationship is limited because both program class students and regular class students are reluctant to get to know each other more closely because of these barriers. Unless they are in the same organization, because they are in the same group they can adapt better.

DISCUSSION

The discussion in this research is based on research data, both the results of statistical analysis and qualitative analysis. The analysis covers several things starting from whether the ability grouping method is in accordance with the initial objectives for which the method was created, how input (students, teachers, facilities and infrastructure) supports the ability grouping method, what is the process of implementing the ability grouping method, what are the student learning outcomes (achievement and moral development) in accordance with the initial objectives of establishing the ability grouping method and finally whether the ability grouping method is effective in the teaching and learning process at MTs. X.

As can be seen in the results of the descriptive analysis, all aspects, both teachers and students, have the largest scores in the good assessment. Even though the student's assessment score is lower than the teacher's. The percentage of good and bad ratings for students only differed slightly, so we continued with the qualitative data results to deepen the reasons. The indicators contained in each aspect are the result of conclusions obtained from interviews with the principal and verified by teachers and students with good and bad assessments obtained from the descriptive results. The following is an explanation of each evaluation aspect based on the results of descriptive analysis.

The first aspect of evaluation is context evaluation. Kellaghan & Stufflebeam (2003) say that context evaluation assesses needs, problems and opportunities in defining the environment. Based on the results of the descriptive analysis above, teachers and students gave a good assessment of this evaluation aspect. The method applied at MTs. X is in accordance with the initial purpose of this method. Based on the results of interviews with students who had the lowest scores on the evaluation scale, the majority of them considered this aspect to be quite good. Starting from the first indicator, namely reviewing the application of classroom methods (program and regular) over several time periods, both students and teachers agree that this method has gone through a long consideration process and involved several parties, in this case teachers and principals in meetings held by the school.

Another indicator is the consideration and purpose of implementing the ability grouping method in MTs. X includes ease in the selection of students for competition events and

becomes a differentiator with other schools and the aim is to print superior students and raise the names of madrasas in general (Islamic symbols). This indicator was also highly approved by both the majority of students and teachers. Based on the results of the interview, students said they agreed with this. However, several program class students who were interviewed did not agree with one of the considerations for implementing this method, namely to make it easier to select students for competition activities. Because for them, all students, both program and regular classes, have their own strengths and talents.

Championships, especially in competitions outside of school or between schools, are actually not all won by program class students. In fact, most championships are won by representatives from school organizations or extracurriculars which consist of program classes and regular classes. The selection of competition events referred to in the aim of applying this method are competition events in the academic field such as science or mathematics Olympiads which will definitely be directly represented by program class students without selection. Apart from that, the majority of students and teachers still agree and it can be concluded that the application of this method is basically in accordance with the initial purpose of implementing this method.

The second aspect is input evaluation. This aspect is used to assess competitive strategies and approaches chosen in work plans and budgets in program implementation (Kellaghan & Stufflebeam, 2003). The input contained in the input evaluation indicators includes the availability of teachers or teaching staff who teach in program and regular classes, the teaching performance of teachers in both classes, the quality of students in both program and regular classes, school facilities and infrastructure, availability of media. learning, as well as other things that support the teaching and learning process such as lesson hours, learning materials and budget or tuition fees. Both teachers and students gave good assessments of several of these aspects. Teachers gave a fairly high assessment with a good percentage of 68%, while the students were only 59%.

The results of interviews with several students who gave poor ratings on this aspect, they agreed that several indicators of this aspect such as lesson hours, learning materials, and teachers who taught both in regular classes and program classes were no different. Likewise with the performance of teachers who teach, there is no difference between regular classes and program classes. It's just that according to program class students, the teacher gives more practice questions in the regular class, while in the program class more material is delivered using the lecture method.

The teachers felt a significant difference. Teachers are sometimes more comfortable teaching in program classes because of their ability to absorb material more quickly, so teachers do not have to expend more energy. Apart from that, some teachers also provide more frequent exams in program classes than regular classes. However, both program and regular classes are recognized by teachers as having their own advantages and disadvantages. If the program class has advantages in academics or related to cognitive aspects, students in regular classes excel in psychomotor aspects.

According to the student affairs department, basically both program classes and regular classes both have their respective strengths and weaknesses. It's just that in terms of learning, teaching in a regular class requires more energy and preparation than teaching in a program class because students' comprehension skills are not as good as program classes. Another indicator that was felt to be less good from the input evaluation aspect felt by the students interviewed due to their low evaluation scores was facilities and infrastructure. The absence of a computer

laboratory at all really hampers the teaching and learning process, especially the implementation of the Computer Based National Examination (UNBK).

This was also felt and acknowledged by the school principal. This program and regular class system is just a separation and grouping of ranked students and non-ranked students without any special program or method given to the program class. The principal said that ideally, program classes would receive different treatment. Likewise in regular classes, teachers who teach in regular classes must be teachers with good quality. So that the process of forming superior students does not only occur in program class students but also regular class students.

The third aspect is process evaluation. Kellaghan & Stufflebeam (2003) added that process evaluation is used to monitor, document and assess activities. Indicators from the evaluation process such as the use of available facilities and infrastructure, teacher teaching motivation and student learning motivation, the system or implementation of the ability grouping method are clear and adequate, teachers who give good ratings are quite high, namely 70%. Meanwhile, only 52% of students gave good ratings. Even though it is greater than the percentage of students who gave unfavorable ratings, the difference is very small. The majority of teachers agree that the facilities and infrastructure have been used optimally, teaching in program classes is more motivating and is considered an achievement, although not all think so.

Several students felt differently who gave the lowest rating on the evaluation scale. Based on the interview results, if you look at the teaching and learning process, basically there is nothing different. They still feel motivated to learn, especially program class students. Because competition is so tight, they are required to continue learning and avoid being left behind by others. However, this does not mean that students in regular classes are not motivated. They also admitted that they had the same motivation to learn. It's just that according to regular class students, they will lose interest in studying or decrease their motivation to learn, if they encounter subjects that they find difficult, such as science and mathematics. Moreover, if this is added to the teacher's unpleasant attitude towards students, students' learning motivation will decrease.

The principal also said that the process of implementing this class system made the teachers who taught in the program classes sometimes lazy to teach. This is because they feel that students in program classes have more abilities than regular classes. It didn't take long for students in the program's classes to understand the material. So the material can be completed more quickly and teachers have more free time.

The final aspect is product evaluation. Product evaluation is used to identify and assess short-term or long-term results, both desirable and undesirable (Kellaghan & Stufflebeam, 2003). Indicators from the product evaluation such as school performance which increased rapidly after the ability grouping method was implemented, KKM for madrasahs which increased, names of MTs. X, which is a private madrasa, is superior in Cipanas District, student enthusiasts who continue to increase every year, the fulfillment of rewards for outstanding students because of the abilities of grouping, the existence of a standard procedure of the System of Based on the Class of Program students who are indicated more individualists.

Based on these indicators, 62% of teachers gave a good assessment and 54% of students gave a good assessment. The majority of teachers and students agree that MTs. X becomes a private MTs who excels in Cipanas District with student enthusiasts who continue to increase every year. The school's achievements continue to increase rapidly. Although in fact, this increase in

achievement is not entirely due to the program and regular class system. Because most championships are won by members of organizations consisting of program and regular students. They took part in the competition without selection. This means that both the program and the regular program in terms of achieving competitive achievements representing outside schools have the same role.

Another finding from this aspect actually came from the FGD process with students who gave low scores on the evaluation scale. Initial indications that students in the program's classes tend to be individualistic are almost evident. This can be seen from the confessions of regular class students who think program class students are always alone and separate themselves from other students. Likewise, program class students are reluctant to join because they feel uncomfortable if they suddenly enter or join regular class students.

Even though program classes and regular classes are no longer separated in different buildings, program classes are in the middle of regular classes, both program and regular students are still reluctant to join and greet each other. Program class students are more seen as separating themselves and are considered exclusive by regular class students. Even though program class students themselves want to be able to join and make friends with regular students. However, most of them feel reluctant because there seems to be a distance or fortress from students in regular classes. This is in line with what Gordon (1991, as quoted in Lie, 2007) said, that basically humans like to gather with those who are the same as them and tend to distance themselves from those who are different.

The principal also added that this also happened to teachers. Many teachers feel proud and more exclusive by teaching in program classes. When instructors carry out a successful program in class they will be very proud. Even though the principal said that teachers should be proud if teaching was more successful in regular classes. All the limitations that regular class students have in terms of academics, only the best and most accomplished teachers should be able to make them superior. So that the process of producing superior students does not only occur in program classes but also in regular classes.

Based on the discussion of the results of the descriptive analysis of each aspect of the evaluation above, there are several advantages and disadvantages of each aspect. Apart from all that, because the final problem formulation of this research is the effectiveness of the ability grouping method on the teaching and learning process, the final discussion is about this effectiveness. Based on the results of hypothesis testing using Z test calculations, the H0 result was rejected, which means that the ability grouping method is effectively used in the teaching and learning process at MTs. X. That is, behind all the deficiencies that exist from the application of this method, things that are expected to exist and achieved in the learning process can be optimized enough so that this method can support the effectiveness of teaching and learning in schools.

The effectiveness of the ability grouping method applied in MTs. X can also be explained through several previous studies which also found the effectiveness of this method. Research conducted by Adodo & Agbayewa (2011) found that grouping students based on ability is a superior way to improve student learning outcomes. Apart from that, these findings are also in line with reports of changes in students' attitudes and interest in the research, which is a positive effect of grouping classes into homogeneous groups. Based on the results of interviews with the school principal, the principal realized that student achievement had increased after implementing this method. It can be seen from the KKM value which has increased rapidly,

approaching 80 with a KKM value of 79. Apart from that, achievements through foreign competition championships have also increased. Although according to student explanations, this achievement was not only given by program class students but also regular students who were members of an organization.

Lie (2007) said that one effective way to make students learn based on their abilities is by grouping. Generally, to overcome this, grouping is done by dividing homogeneous students into the same group based on their achievements, which is a method commonly called the ability grouping method. This method is exactly what is done at MTs. X, by grouping students ranking (homogeneous) with good moral records into a group of program class and students with ranking underneath into the regular class.

Apart from the positive effects of this method, negative results were also found from the FGD process with students. This negative effect occurs on students, especially between program class students and regular class students. Many regular class students feel awkward with students in program classes and vice versa. So they are not free to mingle between program and regular classes, even though they want to do so. Ireson and Hallam (2003) also realized that there were negative effects in implementing this method, especially with this type of streaming. According to him, students may become labeled or branded by teachers or other students, which then allows for dissatisfaction or hostility at school. Although there have been no indications of hostility resulting from this class system, what the program class students feel is that being in this class, the program class students feel more exclusive and the regular class also thinks so. So the regular class becomes awkward if you enter the program class or join the program class students. Class seems to be a barrier to their friendship.

So far, the use of the ability grouping method in MTs. X is still effective. There are only a few things that need to be improved and this is quite in accordance with the statements from the initial interview with the principal. There needs to be clarity on the differences when implementing a program and a regular one. Program classes with their abilities are given different materials and loads from regular classes and vice versa. The teachers who teach will definitely be different. Because teachers who teach in regular classes will be better if the teachers are quite accomplished and able to apply more diverse learning methods so that they can be accepted by students whose learning methods are more diverse in regular classes.

The feeling of more exclusivity that arises from the program class and the regular class also admits that differences arise that limit their friendship, one of which can be overcome by holding a cross-class assignment or project as an additional value where the members are program class students and regular class students who are united. project period until the end of the learning semester. Then the best project is selected which is then presented to all students at class meetings after UAS or other activities. So that both regular class students and program class students can be conditioned to interact with each other with the same position and position, namely as students who hold the same obligations and tasks who must work together with each other.

Another way is like what was obtained in Jembarwati's (2015) research which has proven that one way for students' communication skills to develop more effectively so they can interact with their environment is with an intervention in the form of Cognitive Behavior Modification (CBM). The first stage is to carry out self-observation (listening to the internal dialogue within oneself to recognize the characteristics of one's negative statements, involving activities to increase sensitivity to thoughts, feelings, actions, physiological reactions and reaction patterns

towards other individuals). The second stage is creating a new internal dialogue (thinking about alternatives to deviant behavior or behavior that is the main problem. Students are trained to develop alternative behavior that is adaptive or not deviant by changing the internal dialogue within the individual. At the same time, the individual is expected to remain focused attention to the task of making new statements and observing differences in results before participating in training and after participating in CBM training.

Apart from creating joint projects and CBM training, other techniques, for example, are holding joint activities outside such as study tours, outbound, camps, which are carried out by groups with activities that mix program class and regular class students so that closer togetherness and relationships can be built. better. Activities that require teamwork, good communication between team members, will help students to interact with their group, or other groups.

CONCLUSIONS AND SUGGESTIONS

Based on the results of research and discussion regarding the evaluation of the ability grouping method on the effectiveness of teaching and learning at MTs. X, it was concluded that both teachers and students had better assessment scores than bad ones. So it can be concluded that the application of the ability grouping method is in accordance with the initial objectives of the method.

Descriptive analysis of the input evaluation concluded that both teachers and students had good assessment scores greater than bad ones. So it can be concluded that the input (students, teachers, facilities and infrastructure) is sufficient to support the application of the ability grouping method.

The good assessment given to the process evaluation aspect by both teachers and students, good assessment scores are greater than bad. So it can be concluded that the process of implementing the ability grouping method is considered good by teachers and students.

The final evaluation aspect is product evaluation, where both teachers and students have a good assessment score greater than a bad one. So it can be concluded that student learning outcomes (achievement and moral development) are quite in accordance with the initial objectives of establishing the ability grouping method.

Based on the results of inferential statistical tests, the results showed that the ability grouping method was effectively used in the teaching and learning process for students.
Based on the results of research that has been carried out with all limitations, the researchers put forward several suggestions that can be used as feedback and material for consideration by several parties. First for the school. Based on the results of this research, basically the ability grouping method is quite effective in the teaching and learning process, especially in efforts to group students so that the learning process can be more focused and directed with homogeneous groups of students. It's just that there are several consequences and shortcomings that must be accepted, such as students' social relations between classes (program and regular) which are not very intimate because of the barriers between them. So if the main aim of applying this method is only to make it easier to select students for competition events, without this method this convenience can be achieved.

An example is by mapping students' interests and talents at the beginning of entering new students. Then students with certain talents who will be prepared for certain competition activities are marked and prepared in other forums outside the classroom such as special

tutoring groups without having to be grouped in one class. So that both very smart students and cognitively ordinary students can mix in the same group.

The feeling of more exclusivity that arises from the program class and the regular class also admits that differences arise that limit their friendship, one of which can be overcome by holding a cross-class assignment or project as an additional value where the members are program class students and regular class students who are united. project period until the end of the learning semester. Then the best project is selected which is then presented to all students at class meetings after UAS or other activities. So that both regular class students and program class students can be conditioned to interact with each other with the same position and position, namely as students who hold the same obligations and tasks who must work together with each other.

Another way is like what was obtained in Jembarwati's (2015) research which has proven that one way for students' communication skills to develop more effectively so they can interact with their environment is with an intervention in the form of Cognitive Behavior Modification (CBM). The first stage is to carry out self-observation (listening to the internal dialogue within oneself to recognize the characteristics of one's negative statements, involving activities to increase sensitivity to thoughts, feelings, actions, physiological reactions and reaction patterns towards other individuals). The second stage is creating a new internal dialogue (thinking about alternatives to deviant behavior or behavior that is the main problem. Students are trained to develop alternative behavior that is adaptive or not deviant by changing the internal dialogue within the individual. At the same time, the individual is expected to remain focused attention to the task of making new statements and observing differences in results before participating in training and after participating in CBM training.

Apart from creating joint projects and CBM training, other techniques, for example, are holding joint activities outside such as study tours, outbound, camps, which are carried out by groups with activities that mix program class and regular class students so that closer togetherness and relationships can be built. better. Activities that require teamwork, good communication between team members, will help students to interact with their group, or other groups.

Second is advice for teaching staff (teachers). Teaching staff are expected to continue to develop the learning methods used and adapt these methods to students' abilities. Teachers who teach in regular classes should actually be more creative and innovative in preparing material and teaching methods so that students with cognitive abilities of average or below can understand the material provided. Meanwhile, teachers who teach in program classes can actually be more relaxed in preparing because it doesn't require more effort for students in program classes to better understand the material. They only need to be given a challenge or challenges and rewards in each material to increase their motivation and enthusiasm for learning without having to follow through with full teacher intervention.

The next suggestion is for students. Students are expected to be able to follow the learning process well and optimally according to what has been scheduled and arranged by the school. It would also be good for students to be more active and assertive in expressing their opinions, especially regarding every policy provided by the school. So that the policies provided can also be useful for students.

The last one is for future researchers. This research was conducted at schools that apply the ability grouping method using the streaming type, so it is hoped that future researchers can find

schools that use other types of ability grouping methods so that their effectiveness can be compared. Then, future researchers are expected to be able to add or replace variables other than effectiveness which are thought to be related to evaluation and ability grouping methods, for example intergroup bias, interpersonal conflict and others.

REFERENCES

Adodo, & Agbayewa. (2011). *Effect of Homogenous and Heterogeneous Ability Grouping Class Teaching on Student's Interest, Attitude and Achievement in Integrated Science.* International Journal of Psychology and Counselling, 3(3), 48-54.

Arikunto, S., (2015). *Basics of Educational Evaluation.* Jakarta: PT. Bumi Aksara.

Arikunto, S., & Safruddin, C., (2014). *Evaluation of Educational Programs.* Jakarta: Bumi Literacy.

Azwar, S., (2015). *Preparation of Psychological Scales.* Yogyakarta: Student Library

Bramley, P., (1996). *Evaluating Training Effectiveness.* (R. Bennett, Ed.) England: McGraw-Hill International (UK) Limited.

Clark, G., & Zimmerman, E., (1994). *Programming Opportunities for Gifted and Talented Students in the Virtual Arts.* USA: DIANE Publishing.

Ireson, J., & Halam, S., (2003). *Ability Grouping in Education.* London: SAGE Publications Inc.

Jembarwati, O., (2015). *Behavior Modification for Self-Knowledge in Forming Positive Attitudes of Students.* Psympathic, Scientific Journal of Psychology, 2(1), 57-62.

Kellaghan, T., & Stufflebeam, D., (2003). *International Handbook of Educational Evaluation.* USA: Kluwer Academic Publishers.

Lie, A., (2007). *Cooperative Learning.* Jakarta: PT Grasindo.

Prayitno. (2009). *Basic Theory and Practice of Education.* Jakarta: Grasindo.

Seel, N. M., (2012). *Encyclopedia of the Sciences of Learning.* USA: Springer.

Silalahi, U., (2012). *Social Research Methods.* Bandung: PT. Refika Aditama.

Stufflebeam, D., & Zhang, G., (2017). *The CIPP Evaluation Model.* USA: The Guilford Press.

Sugiyono. (2014). *Combined Research Methods (Mixed Methods).* Bandung: Alphabeta.

Supardi. (2015). *Effective Schools: Basic Concepts and Practice.* Depok: PT. Raja Grafindo Persada.

Hero. (2011). *Evaluation: Theory, Models, Standards, Applications and Professions.* Jakarta: PT. Rajagrafindo Persada.

Yero, J. (2002). *The Meaning of Education.* Retrieved May 3, 2017, from Theacers of Mind: http://www.stoa.org.uk/topics/educations/The%20meaning%20of%20education.pdf.