



A Causal-Comparative Study of Inquiry-Based Science Learning Based on Levels of Students' Cognitive Learning Outcomes: Systematic Review

Bani Nugroho, Zulfiani*

Department of Biology Education, Faculty of Tarbiyah and Teachers' Training, Syarif Hidayatullah State Islamic University Jakarta, Indonesia

*Email: zulfiani@uinjkt.ac.id

DOI: 10.24815/jpsi.v9i4.20579

Article History:

Received: April 4, 2021

Accepted: October 5, 2021

Revised: September 9, 2021

Published: October 11, 2021

Abstract. The effect of inquiry-based learning models on cognitive learning outcomes has been discussed for a long time. However, in local research, this effect is more focused on comparing inquiry-based learning models with traditional learning models. Given such studies' rarity, this study seeks to compare the various inquiry levels to cognitive learning outcomes. This study aims to determine the differences in students' cognitive aspects between structured inquiry and guided inquiry-based science learning models. The research method used is a causal-comparative method, with a sample search technique in the systematic review. The sample used is secondary data in the form of undergraduate theses that have passed the selection and come from the Biology Education program at least accredited B with the research theme of the effect of inquiry learning models on high school students' cognitive learning outcomes. The research findings reveal significant differences in cognitive learning outcomes between the structured inquiry (SI) and guided inquiry (GI) learning model. The processing is more complicated in the GI learning model, allows students to perform better in learning outcomes than the SI learning model. Significant differences were supported by calculating the effect size in this study. The effect size in studies that apply the SI learning model belongs to the medium category. Meanwhile, the effect size in studies using the GI learning model belongs to the large to extremely large categories.

Keywords: cognitive, guided inquiry, structured inquiry.

Introduction

Inspired by teacher learning engineering and student learning actions in the book Dimiyati and Mudjiono (2006) entitled *Learn and Learning*, wherein achieving a learning outcome, the teacher has a role in compiling instructional designs to teach students. To carry out these teaching and learning activities, teachers need to compose or develop a science learning process. Learning science in schools expects to stimulate students to find their problems and solutions to solve these problems through their ability to think independently. In line with the new paradigm in science learning, student learning is encouraged to learn more facts and concepts verbally through memorization. Nevertheless, educators should be more active in providing students' experiences to understand better, then guide students to develop this knowledge in real life (Puspita, et al., 2018).

Unfortunately, during obtaining facts, concepts, and theories in schools' science learning process, it is still centered on the teacher by conveying them directly rather than students looking for them independently. There are still learning models applied in schools that use traditional learning models, where the teacher dominates the active learning process. As a result, students are increasingly passive in conveying ideas or ideas and are

easily bored when teaching and learning activities occur. The learning outcomes that students will get are not optimal (Khalaf & Zin, 2018). Based on these problems, an effort is needed to be able to optimize student learning outcomes (Yasniati, 2017). Among them, namely by applying an inquiry-based learning model.

Inquiry-based learning or discovery can direct and lead students to systematically solve a problem by providing students opportunities to be free to carry out their experiments to become a sustainable and suitable solution to be carried out. This can reflect the student's personality, reflected in everyday life in finding answers or solutions to problems using existing clues. This is part of the learning outcomes based on inquiry itself. Indirectly, affective and psychomotor competencies and students' independent character are formed from getting used to finding solutions through investigation.

Inquiry-based learning stages generally include the ability to formulate problems, design experiments, collect data, and draw conclusions. Through this inquiry process, students become involved in realizing meaningful, creative, critical, and problem-solving learning. This learning process focuses on several aspects. Among others the development of cognitive, psychomotor, affective, and metacognition (Zulfiani & Herlanti, 2018; Ulfah, et al., 2021).

Inquiry-based learning models have been known to have various levels. When students experience these various inquiry levels, they will develop the ability and understanding of scientific inquiry. Learners need to experience science through hands-on experience, consistently practice inquiry skills and seek a deeper understanding of science's content through their investigations. Achieving these goals is feasible after the teacher can identify the level of inquiry in science material and revise it as needed to accept the various complexities in their inquiry experiences. Based on these differences, it is necessary to know the impact on student learning outcomes (Zulfiani, et al., 2018).

Preliminary studies are carried out by looking for research that has been done before. The research that was collected was research on inquiry-based science learning. Was obtained temporarily with a range of 2013-2019 as many as eight research titles conducted by Biology *Tadris* Students Faculty of Tarbiyah and Teacher Training at UIN Syarif Hidayatullah Jakarta, both types of quasi-experimental research and comparative and correlation research (UIN Syarif Hidayatullah Institutional Repository, 2020). With this research, it is interesting to accumulate in a study to see the differences in inquiry-based science learning (structured and guided) on student learning outcomes. However, many undergraduate thesis research of Biology Education students at other universities, both published online and unpublished regarding inquiry-based learning, is necessary to organize data and systematic review of the data obtained and afterward.

Methods

This study uses a comparative causal research method with a sample search technique in a systematic review (Aktamis, et al., 2016). The time of implementation of this research is from July 2020 to December 2020. This research's population or object is the Undergraduate Thesis with the theme of inquiry-based science learning (structured inquiry models and guided inquiry) on student learning outcomes from study program accredited A or B. In comparison, the research subjects are high school students. The samples in this study were undergraduate theses selected based on predetermined inclusion and exclusion criteria. The instrument used is a form in the form of a table of the selection, coding, assessment of the study's quality, the Google Scholar search engine, and institutional repositories.

The combination of the Google Scholar search engine and the institutional repositories is considered based on the sample selected in this study, namely in an undergraduate thesis. According to Bramer et al. (2017), when looking for references or relevant research

samples in systematic reviews, it is recommended to use multiple databases. In his research, it was found that the combination of Embase, MEDLINE, Web of Science Core Collection, and Google Scholar performed best as a search engine for systematic reviews. Gusenbauer & Haddadway (2019), in their research, found that there are 14 academic search systems, including Goog

le Scholar are categorized as supplementary search systems. While the repository was chosen as a digital database in this study, it was strengthened by Suwanto (2017) research. Where the existence of an internet network and the discovery of a software repository, scientific papers compiled by students or other academics, which initially could only be accessed on a limited basis area and place), becomes infinite.

Referring to the Higher Education Database (PDDikti) page of the Ministry of Education and Culture of the Republic of Indonesia, 110 universities have a Biology Education or *Tadris* study program minimum B accredited (Ministry of Education and Culture of the Republic of Indonesia, 2020). Of the 110 tertiary institutions, only 87 have access to their digital library. Furthermore, from the 87 repositories, about 25% only provide the student thesis file in the published version.

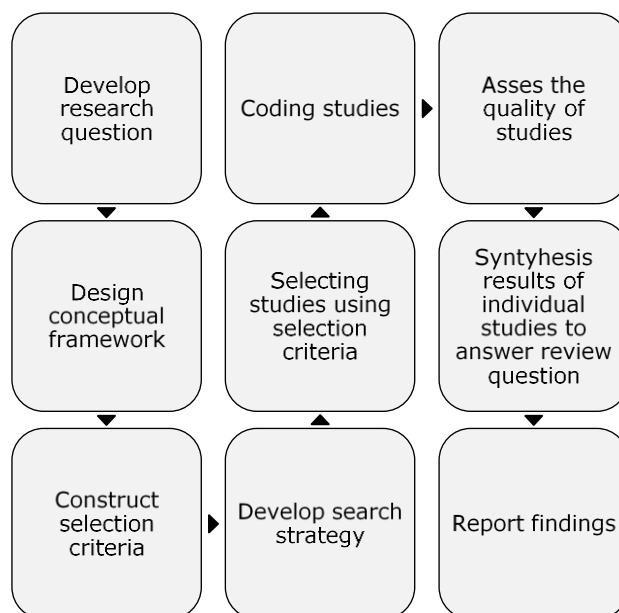


Figure 1. The Research Process.

This research's stages were adapted from Richter (2020), which is presented in Figure 1. The inclusion criteria in this study, among others: 1). Undergraduate Thesis from Biology Education deals with the influence of inquiry-based science learning (structured inquiry models and guided inquiry) on student learning outcomes, seen based on the title or thesis abstract; 2). The Honorary Board has approved the faculty's undergraduate thesis concerned marked with a signature and stamp on the thesis or thesis validation sheet available (has been uploaded) in the official digital library of the author's university of origin; 3). Thesis passed 2010-2019; 4). An undergraduate thesis can be accessed in the form of a published version; 5). The undergraduate thesis research sample is students at the high school level (SHS). While the exclusion criteria in this study, among others: 1). Undergraduate thesis comes from a Biology Education study program that has not been accredited or accredited under B; 2). There is no data on learning outcomes in the undergraduate thesis in pre-test/post-test scores or the average research score of each experimental/control group (thesis attachment).

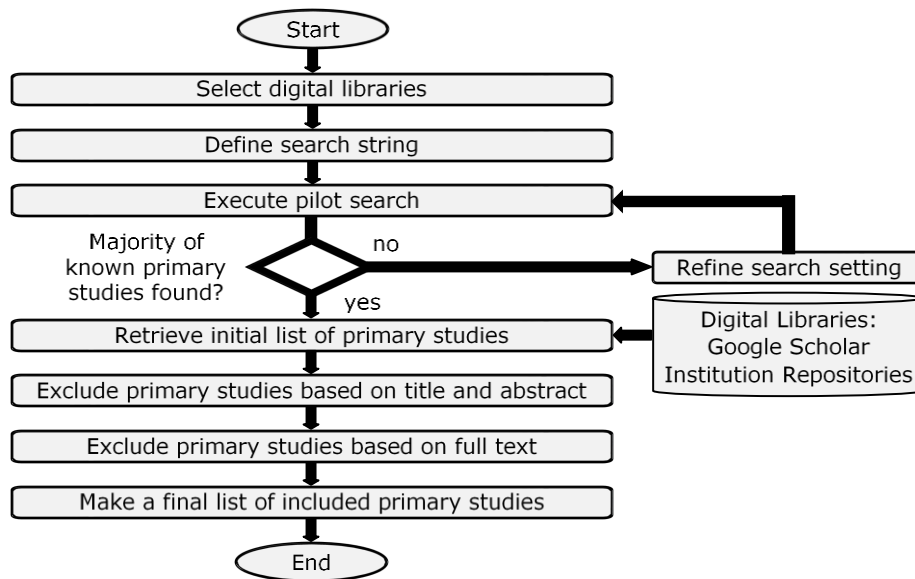


Figure 2. Search and Selection of Primary Studies.

The systematic review requires thorough, objective, and reproducible searches from multiple sources. A systematic review is useful for identifying as many eligible studies as possible (within resource limits) to minimize bias and achieve more reliable estimates of effects and uncertainties (Higgins, et al., 2019). The detailed search stages in this research are present in Figure 2. Among them are determining digital library sources, creating search strings, carrying out search experiments, refining search strings, and taking the initial list of primary studies that match digital library sources' search strings (Wahono, 2015). The search strings used in this study are as follows: (*inkuiri* OR inquiry) AND (*terstruktur* OR structured) AND (*terbimbing* OR guided) AND (*kognitif* OR cognitive) AND (*skripsi* OR thesis) NOT (*jurnal* OR journal).

This research's data analysis technique is through a process flow adapted from the Statistics book for Educational Researchers by Lolombulan (2017). The analysis used is the Kruskal Wallis test and the Bonferroni continued test based on the data obtained. To get a good analysis report, apart report the significance value, Lee (2016) recommended including the effect size value so that the research report is more substantial. Even recent research guidelines recommend that researchers also report effect sizes for the interventions or associations studied (Berben, et al., 2012). The effect size estimate is a measure worth reporting alongside the p-value in testing the null hypothesis. The effect size can be used in any type of quantitative research to show the size of the influence of a variable or study as a whole (Furtak et al., 2012; Tomczak, M. & Tomczak, E., 2014). Here are two effect size formulas used in this study (Carlson & Schmidt, 1999).

$$ES_{SGPP} = d = \frac{(T_2 - T_1)}{S_{T_1}} \quad (1)$$

$$ES_{POWC} = d = \frac{(T_2 - C_2)}{S_{T_2 C_2}} = \frac{(T_2 - C_2)}{\sqrt{\frac{(S_{T_2})^2 + (S_{C_2})^2}{2}}} \quad (2)$$

Remark:

- ES_{SGPP} : Effect Size for Single Group Pre-Test/Post Test Design
- ES_{POWC} : Effect Size for Post Test Only with Control Group Design
- T_1 : Mean Pre-Test of Treatment Group

- T_2 : Mean Post Test of Treatment Group
- C_2 : Mean Post Test of Control Group
- S_{T_1} : Standard Deviation of Pre-Test Treatment Group
- $S_{T_2C_2}$: Standard Deviation of Pooled Estimated

Table 1. Labels for Values of d.

Value of d	Label
0,20	Small
0,50	Medium
0,80	Large
1,10	Very Large
1,40 +	Extremely Large

Table 1 explains the meaning of the effect size (d). The value of d obtained is in the range of 0.20 to less than 0.50, the effect size is classified into the small category. Furthermore, if d obtained's value is equal from 0.50 to below 0.80, it is in the medium category. Meanwhile, if the d obtained value ranges from 0.80 to below 1.10, it is a large category. If the value of d obtained is 1.10 to below 1.40, it belongs to a very large category. Finally, if the d value obtained is equal to 1.40, it can be said that it belongs to the extremely large category (Patten & Newhart, 2017).

Results and Discussion

Based on systematic literature review, it is known that quite a lot of local empirical literature compares inquiry-based learning models with non-inquiry-based learning models on learning outcomes. However, few studies have focused on the differences between the various levels of inquiry. Given such studies' rarity, this study seeks to compare the various inquiry levels to cognitive learning outcomes. However, what can be realized in this study is only to compare the two levels of the inquiry learning model, namely structured inquiry and guided inquiry. The two levels differ in explicit instructions (for example, guide questions, procedures, and solutions or expected results). The following describes the discussions on the findings of this study.

Table 2. List of Included Studies and Codes Applied.

Researchers (Year)	Topics	Class	Study Design	Inquiry Levels	Respondents	Study Codes
Kusumawati (2017)	Ecosystem	X	Single group pre-test/post test	1	40	A 1
Megawati (2018)	Protists	X	Single group pre-test/post test	2	30	B 2
Lestari (2019)	Ecosystem	X	Post test only with control group	2	33	C 2
Muslimah (2019)	Fungi	X	Single group pre-test/post test	2	36	D 2
Setiawati (2018)	Animalia	X	Single group pre-test/post test	2	40	E 2

Based on the five studies or samples in Table 2 that have successfully passed the previous selection stages, all of them are given a study code with the alphabet A - E in sequence, followed by an inquiry level code with numbers 1-2. Based on the stages of coding the study names that have been carried out in this study, the study codes are A 1, B 2, C 2, D 2, and E 2. Study A 1 means study A with the level of inquiry category 1, namely structured inquiry. Study B 2 means study B with a category 2 level of inquiry, namely guided inquiry, and so on.

Petticrew and Roberts (2005) in Richter (2020) mention that in the context of a systematic review, assessment of study quality is often referred to as 'Critical Assessment. According to Gough (2007) in Richter (2020), three elements are widely considered in the critical assessment: the suitability of the study design; the quality of the implementation of the study method; and the relevance of the study to the review questions. In this study, study quality research was applied to each sample to review aspects of the implementation of the research methodology. Based on the assessment of the quality of the study on all samples whose research methods are both quasi-experimental, several strengths and weaknesses of each were obtained. From the assessment of the quality of the study, the decision is that all studies are worthy of further analysis with note-taking into account the strengths and weaknesses that exist.

Referring to five studies that have passed the study quality assessment stage and have been declared feasible, the next step is to analyze all these samples' data. The research data used is raw data that comes from the appendix of each study. Furthermore, the data were analyzed descriptively with IBM SPSS Statistics 26 software to verify the calculations' accuracy and report on each study. Based on this re-analysis (evaluation), it shows the suitability or accuracy of the statistical descriptions with the analyzes that have been previously reported in each study.

Table 3. Statistical Description of Learning Outcomes Data (Post Test) for the Experiment Class of 5 Studies

		Study Codes				
		A 1	B 2	C 2	D 2	E 2
N		40	30	33	36	40
Mean		70,58	77,50	80,52	83,08	80,13
95% Confidence Interval for Mean	Lower Bound	68,18	74,26	78,73	81,73	76,94
	Upper Bound	72,97	80,74	82,30	84,44	83,31
Median		71,00	77,50	81,00	83,00	80,00
Std. Deviation		7,50	8,69	5,04	4,01	9,97
Minimum		58	65	71	73	60
Maximum		88	90	93	87	100

Table 3 shows the comparison of the statistical descriptions of learning outcomes from the five studies that were the research samples. The table shows that each study's total respondents ranged from 30 to 40 students (categories of less than 50 respondents). Next, each study's mean learning outcomes listed in Table 3 show the diversity of values. The most important mean of 83.08 is owned by study D 2, and the smallest is 70.58 owned by study A 1. The lower limit and the upper limit of study A1 at the 95% confidence interval for mean are the smallest among other studies, namely 68.18 and 72.97. Simultaneously, the lower limit and the most prominent upper limit are owned by the D 2 study, namely 81.73 and 84.44. Then it is the same as the previous position, where the smallest mean

value is owned by Study A1, which is 71.00. Meanwhile, the enormous mean value is owned by the D 2 study, which is 83.00.

Based on Table 3, the distance between data points with an average value for the largest is owned by Study E 2. This is indicated by the enormous standard deviation value of study E 2 compared to other studies, 9.97. On the other hand, the distance between data points with the smallest mean value is owned by the study code D 2 with a standard deviation of 4.01. Next, Table 3 presents the minimum and maximum value of learning outcomes that each study has. The variation in the five studies' minimum value ranges from 58 to 71. The maximum value variation is from 87 to 100, where the minimum value is the smallest in Study A2, which is 58.00. The most incredible maximum value is owned by study E 2, which is equal to 100.

Table 4. Normality Test

	Study Codes	Shapiro-Wilk Sig.
Cognitive Learning Outcomes	A 1	0,08
	B 2	0,02
	C 2	0,33
	D 2	0,00
	E 2	0,23

Testing Requirements Analysis and Hypothesis Testing. Referring to the data analysis techniques that have been determined in Methods, in order to carry out hypothesis testing in this study, it is necessary to carry out a classical assumption test (normality test) and a variance homogeneity test. Based on the normality test carried out by the Shapiro-Wilk test (presented in Table 4), a significance value was obtained greater than 0.05 in studies A 1, C 2, and E 2. This shows that the research data in these studies are typically distributed. Meanwhile, for B2 and D2 studies, the significance value is less than 0.05, which means that these studies' research data are not normally distributed.

Due to the presence of study data that is not normally distributed, namely in the B 2 and D 2 studies, it is necessary to transform the data in each study as a solution so that the data can be normally distributed (George & Mallery, 2020). Data transformation involves applying mathematical procedures to data to appear more normal (Mertler & Reinhart, 2017). In transforming data, the first step that needs to be considered is to look at the slope's direction (skewness).

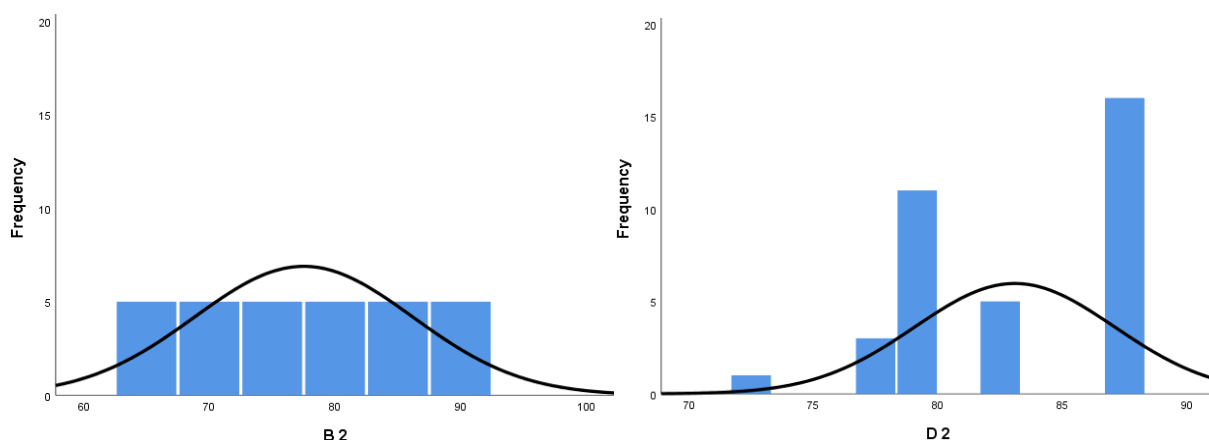


Figure 4. Histograms of Studies B 2 and D 2 (Not in a Normal Distribution)

Figure 4 shows that the data in the B 2 study tilted weakly towards the right (positive), so-called moderate positive skew. Whereas in the D 2 study, the slope tended to flare weakly to the left (negative), usually called a moderate negative skew. Therefore, there are two types of data transformations used in each study. The first transformation is the square root type (Sqrt), referring to study B 2. The second alternative transformation is the square root type reflected in the D 2 study (Tabachnick & Fidell, 2019). After the first type of data transformation and the second type of transformation were carried out in each study, the normality test was again carried out to ensure that the data to be tested for the hypothesis were usually distributed. After that, it can be determined the parametric or non-parametric hypothesis test will be used.

Table 5. Normality Test After Transformation

Study Codes	Shapiro-Wilk Sig.	
	Square Root Transformation	Reflect + Square Root Transformation
A 1	0,07	0.07
B 2	0,01	0.01
C 2	0,40	0.07
D 2	0,00	0.00
E 2	0,20	0.08

Table 5 shows the normality test results after all studies' data were transformed into square root types. Based on the table, it can be seen that the B 2 and D 2 studies are still not generally distributed because the significance value for the Shapiro-Wilk test is less than 0.05. So it is necessary to carry out the second type of transformation, namely the transformation of the reflected square root.

Based on Table 5, the normality test results after the data in all studies were carried out the second type of transformation, namely the reflected square root. The results were the same as the previous one, wherein in study B 2 and study D 2, it was stated that they were not normally distributed. Study B 2 and Study D 2, the Shapiro-Wilk test's significance value, is less than 0.05.

Based on the normality test that has been carried out both before and after the transformation, it confirms that the data in one study were not normally distributed. So there is no need to do the homogeneity of variance test. Therefore, the hypothesis test in this study cannot be done parametrically.

Table 6. Hypothesis Test (Kruskal-Wallis) Summary

Null Hypothesis	Test	Sig.	Decision
The distribution of Learning Outcomes (Cognitive) is the same across categories of Inquiry Learning Models.	Independent-Samples Kruskal-Wallis Test	0,00	Reject the null hypothesis

Since this research cannot be done parametrically because the assumption test results state that the data is not normally distributed, the non-parametric test, namely the Kruskal-Wallis test. A summary of the Kruskal-Wallis test is shown in Table 6, where the P-value is $0.00 < 0.05$. So the decision is to reject the null hypothesis, which says "the distribution of learning outcomes (cognitive) is the same among the categories of inquiry learning models." So there is robust evidence showing differences in at least one group pair.

Research Findings. The decision to test the hypothesis in this study is to reject the null hypothesis, which reads: "The distribution of learning outcomes (cognitive) is the same between categories of inquiry learning models." This suggests that there are differences in learning outcomes between the studies studied. A pairwise comparisons test is

required to determine which study pairs differ, namely the Bonferroni test (Harlan, 2018). The Bonferroni test results are presented in Table 7, and it can be seen that the paired Bonferroni correction value for each sample.

Table 7. Pairwise Comparisons of Learning Outcomes from Included Studies.

No	Sample 1-Sample 2	Adj. Sig.	No	Sample 1-Sample 2	Adj. Sig.
1.	A 1 - B 2	0,02	6.	B 2 - D 2	0,03
2.	A 1 - C 2	0,00	7.	B 2 - E 2	1,00
3.	A 1 - D 2	0,00	8.	C 2 - D 2	1,00
4.	A 1 - E 2	0,00	9.	C 2 - E 2	1,00
5.	B 2 - C 2	1,00	10.	D 2 - E 2	0,50

Based on the Bonferroni correction values presented in Table 7, it can be seen that study A1 paired with other studies shows several 0.00 (except for comparison to Study B2 of 0.02). This concludes that: "Study A 1, namely the structured inquiry (SI) learning model, is significantly different from the other four studies, namely studies B 2, C 2, D 2, and E 2 on learning outcomes (cognitive). It has been known at the outset that the B 2, C 2, D 2, and E 2 studies both use the guided inquiry (GI) learning model.

Among the studies that apply the guided inquiry learning model, almost all fail to reject the null hypothesis. In other words, learning outcomes (cognitive) are the same for studies using GI. This is indicated by the Bonferroni correction value of 1.00 in the comparison of studies B 2 - C 2, B 2 - E 2, C 2 - D 2, and C 2 - E 2 and a Bonferroni correction value of 0.50 in the comparison of studies D 2 - E 2. Only the comparisons for the B 2 - D 2 studies show slight differences from the GI studies. A visualization of the Bonferroni test findings is presented in Figure 5 in the following graphical form.

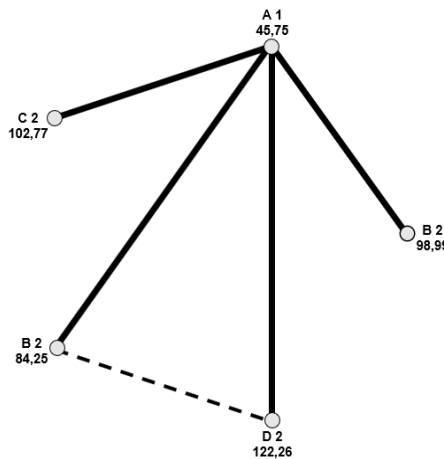


Figure 5. Pairwise Comparisons Learning Outcomes Chart of Included Studies.

Each small circle in Figure 5 shows the mean ranking in each study. Based on Figure 5, it is clear that there are four thick lines. The thick lines appear to both be sourced from or point to Study A 1. These lines represent significant differences between couples. Besides, there is also a dashed line connecting Study B 2 with study D 2, where the dotted line indicates a difference but is less significant.

Students who are taught using the guided inquiry learning model show significant differences in learning outcomes compared to students who are taught using a structured inquiry learning model. This is shown based on the significance test of the experimental group's post-test scores in each study. The Bonferroni follow-up test results state significant differences in learning outcomes between students who are taught using a structured inquiry learning model with a guided inquiry learning model. The post-test value

testing in the experimental group from each study sampled in this study was considered based on a comparative causal research design that examines differences in learning outcomes from two levels of the inquiry learning model, namely SI and GI.

The researchers' alternative to avoid bias regarding respondents' equality was by matching the respondents between studies as suggested by Mills & Gay (2019). Matching that can be done compares the pre-test value and the gain score, and the normalized gain score in each study. The following is an equivalent based on the gain score and N-Gain score between studies.

Table 8. Matching-based Gain and Normalized Gain Score of Included Studies.

Study Codes	Respondents	Mean Pre-Test	Mean Post-Test	Gain Score	N-Gain Score	N-Gain Label
A 1	40	64,70	70,58	5,88	0.10	Low
B 2	30	57,50	77,50	20,00	0.46	Medium
C 2	33	Low	80,52			
D 2	36	50,53	83,08	32,56	0.63	Medium
E 2	40	37,38	80,13	42,70	0.70	High

Based on Table 8, a comparison between studies can be made. Where in the matching stage, the pre-test scores reported in studies A1, B 2, D 2, and E 2 show that the scores are low, which is less than 70. For the C 2 study, the pre-test scores were reported qualitatively, where it was stated that the results of student learning are also classified as low (Lestari, 2019). Furthermore, the Gain score (the difference between the post-test and pre-test) and the Normalized Gain (N-Gain) can be calculated from each study (except for the C 2 study, which was a post-test only design). The amount of Gain and Normalized Gain in learning using the guided inquiry (GI) learning model is much higher than the structured inquiry (SI) learning model.

It is known that study A1 which applies the SI learning model, has a relatively higher pre-test mean when compared to other studies that use the GI learning model. So it can be said that students who use the GI learning model have a higher learning outcome than the SI learning model. The calculation of N-Gain study A 1, namely using SI, is classified as low (even very low) because it is less than 0.3. Whereas for the B 2, D 2, and E 2 studies that use GI, they are classified as moderate categories referring to the N-Gain category created by Hake (1999).

There are similarities in study A 1 with study C 2 in terms of the material concept used. Namely, they both use the material concept of the Ecosystem. It is known that the Bonferroni paired follow-up test showed significant differences between study A1 which used a structured inquiry-based learning model and the other four studies that used a guided inquiry learning model. One of them is study C 2. The significant difference between the two studies strengthens the findings of this study. The similarity of the concept of material between the two studies still shows differences in improving students' cognitive learning outcomes. So it can be said that the guided inquiry-based learning model is better than structured inquiry in terms of improving the learning outcomes of the cognitive aspects of students for the high school (SMA) level.

Referring to the syntax in the Appendix of the Learning Implementation Plan (RPP) for the material concept of the Ecosystem in Study A 1 and Study C 2, it can be seen that there are striking differences in the stages of the process/work steps. In study A 1, the teacher gives a Student Worksheet (LKPD), fully presenting the work steps. Whereas in study C 2, the LKPD provided by the teacher did not provide complete work steps, but students were guided to develop their work steps through a literature review. The stages contained in the guided inquiry learning model include exploration of phenomena, focusing questions, planning investigations, conducting investigations, analyzing investigation

results, constructing new knowledge, and communicating the acquired knowledge ((Baharuddin, et al., 2017; Mauritha, et al., 2017; Sari, et al., 2020). So that this is the main factor that students in the guided inquiry-based learning model are more active and independent than those using the structured inquiry learning model.

Lestari (2019), in his research, namely the C2 study, stated that students who during the learning process use the guided inquiry learning model are required to be able to express their initial knowledge or orientation stage, formulate problems, propose hypotheses, collect data, test hypotheses, and make conclusions. Meanwhile, Kusumawati (2017), in her research, namely in study A1, stated that during the learning process, students were still directed in terms of initial knowledge or the orientation stage to the presentation of procedures/work steps to collect data. Based on the two studies, it can be seen that in the guided inquiry-based learning model, students will be more required to be independent since the process/work steps to collect data compared to structured inquiry.

Table 9. Effect Size Calculation of Include Studies.

Study Codes	Respondents	Treatment		Control	Standard Deviation	Effect Size (d)	Label
		Mean Pre-Test	Mean Post-Test	Mean Post-Test			
A 1	40	64,70	70,58		9,88 ^(Pre-Test)	0,60	Medium
B 2	30	57,50	77,50		8,69 ^(Pre-Test)	2,30	Extremely Large
C 2	33		80,52	74,93	4,89 ^(Pooled)	1,14	Very Large
D 2	36	50,53	83,08		11,97 ^(Pre-Test)	2,72	Extremely Large
E 2	40	37,38	80,13		9,74 ^(Pre-Test)	4,39	Extremely Large

The effect size calculation results presented in Table 9 are used as reinforcement of previous significance tests. Based on Table 9, it can be seen that only study A1 has the effect size value belonging to the medium category. Whereas in the C 2 study, the effect size was classified into a very large category. The effect sizes for studies B 2, D 2, and E 2 fall into the extremely large category. This indicates that the GI learning model has a more significant influence on learning outcomes when compared to the SI learning model. Thus, the calculation of the effect size in this study is in line with the study conducted by Batdi et al. (2018), where 27 studies showed that all levels of the inquiry-based learning model were quite influential in improving student learning outcomes. Research by Lazonder & Harmsen (2016) reveals that the overall average effect size of these 60 studies based on the level of inquiry has a significant positive effect on learning outcomes. Based on the analysis of Abdurrahman's research, it is stated that almost of the study showed that inquiry-based learning has shown students' increasing in cognitive achievement significantly (Abdurrahman, 2017).

The effect size calculation in this study uses the effect size formula, which interprets one variable: the inquiry learning model's effect applied to each study's experimental class on learning outcomes (Carlson & Schmidt, 1999). Therefore, it is not the whole study as a whole. This research report does not end up with a meta-analysis that incidentally looks for the study's effects as a whole. This is considered because it focuses solely on this study's independent variable, namely the inquiry learning model level.

According to Bunterm et al. (2014), students who are faced with a higher form of inquiry or, in other meanings, given less explicit information, in this case, guided inquiry,

show a more significant increase in learning outcomes than students who are taught using structured inquiry. Compared to a structured inquiry, the guided inquiry state was engineered to include more flexibility in the teacher's type and amount of information. Although sometimes, teachers provide more information in the guided inquiry than in structured inquiry conditions, the information is more contextualized with learners' uncertainties. Students are also encouraged to engage, think about, and explain the phenomena they observe (Bunterm, et al., 2014).

Craik & Tulving (1975) argued that more painstaking processing leads to better retention of information. In Bunterm et al. (2014) research report, it was explained that students in the GI group had to make their procedures and analyze the experiment. On the other hand, learners in the SI group received explicit instruction on experimenting and accessing information from the students' textbooks and worksheets. So it can be said, students in a condition of guided inquiry must be involved with more profound information. Perhaps this more complex type of processing allows students to perform better in learning outcomes (Bunterm, et al., 2014). It should be noted that students in the SI group still showed a significant increase, which means that the SI learning model is still classified as effective in improving learning outcomes.

In line with this study's findings, Sadeh & Zion (2009) stated that GI is more effective than SI in conveying science content and science process skills. The same is reported in the local literature. Fahrurrizal et al. (2019) states that learning using the Socio Scientific Inquiry (SSIQ), Guided Inquiry (GI), and Structured Inquiry (SI) model can significantly improve cognitive abilities in high school students. The SSIQ and GI learning model is more effective than the SI learning model (Fahrurrizal, et al., 2019). The GI learning model or guided inquiry is the right choice in providing access to students to fully explore their knowledge actively and independently (Zarisa & Saminan, 2017; Zulfiani, et al., 2018).

The comparative effect of guided inquiry learning models and structured inquiry on student achievement or learning outcomes in Basic Science and Technology was also analyzed by Audu et al. (2017). They revealed that students who were taught using a guided inquiry learning model had higher learning outcomes than students who were taught using a structured inquiry learning model. Based on their research, this may be related to the fact that guided inquiry actively involves students and allows cooperative group participation. It also helps students have in-depth knowledge that is more meaningful when compared to structured inquiry. In this study, students in the GI group had to create their procedures and analyze their experiments.

Meanwhile, students in the SI group received explicit instruction on conducting experiments and had access to information from practicum modules and textbooks. So that students in the GI group must be involved with more profound information. So it is more challenging to process like this that allows them to perform better in terms of achieving learning outcomes (Audu, et al., 2017).

Conclusion

Based on the research that has been done, it can be concluded that there are significant differences in cognitive learning outcomes between the structured inquiry (SI) learning model and guided inquiry (GI). The more complex processing in the GI learning model allows students to perform better in cognitive learning outcomes when compared to the SI learning model. This is supported by calculating the effect size in this study, where the effect size in studies that apply the SI learning model is classified into the medium category. Meanwhile, the effect size in studies that apply the GI learning model is classified into very large to extremely large categories.

Acknowledgment

This research was implemented in good cooperation, and the author expresses gratitude to the various parties who have helped and supported the authors to complete this research.

References

- Abdurrahman. 2017. Efektivitas dan kendala pembelajaran sains berbasis inkuiri terhadap capaian dimensi kognitif siswa: meta analisis. *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah*, 02(1):1-9. <http://dx.doi.org/10.24042/tadris.v2i1.1206>
- Aktamis, H., Hiğde, E., & Özden, B. 2016. Effects of the inquiry-based learning method on students' achievement, science process skills and attitudes towards science: a meta-analysis science. *Journal of Turkish Science Education*, 13(4):248-261. <http://dx.doi.org/10.12973/tused.10183a>
- Audu, C., Ajayi, V.O., & Angura, T. 2017. Do guided and structured inquiry instructional strategies have any comparative effects on students' achievement in basic science and technology? A field report. *Journal of Education and Practice*, 8(33):81-88. <http://dx.doi.org/10.2139/ssrn.3069288>
- Baharuddin, Indana, S., & Koestiari, T. 2017. Perangkat pembelajaran IPA berbasis inkuiri terbimbing dengan tugas proyek materi sistem ekskresi untuk menuntaskan hasil belajar siswa SMP. *Jurnal IPA dan Pembelajaran IPA*, 1(1):81-97.
- Batdi, V., Semerci C., & Aslan, A. 2018. A meta-analytic and thematic study concerning the effect of inquiry based instruction on learners' achievement. *Educational Policy Analysis and Strategic Research*, 13(2):51-68. <http://dx.doi.org/10.29329/epasr.2018.143.3>
- Berben, L., Sereika, S.M., & Engberg, S. 2012. Effect size estimation: methods and examples. *International Journal of Nursing Studies*, 49(8):1039-1047. <http://dx.doi.org/10.1016/j.ijnurstu.2012.01.015>
- Bramer, W.M., Rethlefsen, M.L., Kleijnen, J., & Franco, O.H. 2017. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Systematic Reviews*, 6(245):1-12. <http://dx.doi.org/10.1186/s13643-017-0644-y>
- Bunterm, T., Lee, K., Kong, J.N.L., Srikoon, S., Vangpoomyai, P., Rattanaavongsa, J., & Rachahoon, G. 2014. Do different levels of inquiry lead to different learning outcomes? A comparison between guided and structured inquiry. *International Journal of Science Education*, 36(12):1937-1959. <http://dx.doi.org/10.1080/09500693.2014.886347>
- Carlson, K.D. & Schmidt, F.L. 1999. Impact of experimental design on effect size: findings from the research literature on training. *Journal of Applied Psychology*, 84(6):851-862. <https://doi.org/10.1037/0021-9010.84.6.851>
- Craik, F.I.M. & Tulving, E. 1975. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3):268-294. <http://dx.doi.org/10.1037/0096-3445.104.3.268>

- Fahrurrizal, M., Suwono, H., Susilo, H. 2019. Studi komparatif strategi pembelajaran inkuiri ditinjau dari kemampuan kognitif siswa SMA. *Jurnal Pendidikan*, 4(6):753-757. <http://dx.doi.org/10.17977/jptpp.v4i6.12497>
- Furtak, E.M., Seidel, T., Iverson, H., & Briggs, D.C. 2012. Experimental and quasi-experimental studies of inquiry-based science teaching: a metaanalysis. *Review of Educational Research*, 82(3):300-329. <http://dx.doi.org/10.3102/0034654312457206>
- George, D. & Mallery, P. 2020. *IBM SPSS statistics 26 step by step a simple guide and reference*, 16th edition, Routledge Taylor & Francis, New York.
- Gusenbauer, M., & Haddaway, N.R. 2020. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of google scholar, pubmed, and 26 other resources. *Res Syn Meth*, 11(2):181-217. <http://dx.doi.org/10.1002/jrsm.1378>
- Hake, R.R. 1999. Analyzing change/gain scores. *American Educational Research Association's Division D*, 1(1):1-4.
- Harlan, J. 2018. *Analisis variansi*, Gunadarma, Depok.
- Higgins, J.P.T., Thomas, J., Chandler, J., Chumpston, M., Li, T., & Page, M.J. 2019. *Cochrane handbook for systematic reviews of interventions*, 2nd edition, Welch, V.A. (eds.), John Wiley & Sons, Chichester, UK.
- Khalaf, B.K. & Zin, Z.B.M. 2018. Traditional and inquiry-based learning pedagogy: a systematic critical review. *International Journal of Instruction*, 11(4): 545-564. <http://dx.doi.org/10.12973/IJI.2018.11434A>
- Kementerian Pendidikan dan Kebudayaan Republik Indonesia. 2020. Pangkalan Data Pendidikan Tinggi (PDDIKTI) 2020, daftar program studi Pendidikan Biologi. https://pddikti.kemdikbud.go.id/search_prodi/Pendidikan%20Biologi/-/-/-/S1
- Kementerian Pendidikan dan Kebudayaan Republik Indonesia. 2020. Pangkalan Data Pendidikan Tinggi (PDDIKTI) 2020, daftar program studi Tadris Biologi. https://pddikti.kemdikbud.go.id/search_prodi/Tadris%20Biologi/-/-/-/S1
- Kusumawati, R.I. 2017. Efektifitas pembelajaran inkuiri terstruktur terhadap hasil belajar kognitif, afektif dan psikomotorik siswa pada materi ekosistem kelas X SMA Negeri 1 Comal tahun ajaran 2015/2016. *Undergraduate Thesis*, Universitas PGRI Semarang, Indonesia.
- Lazonder, A.W. & Harmsen, R. 2016. Meta-analysis of inquiry-based learning: effects of guidance. *Review of Educational Research*, 86(3):681-718. <http://dx.doi.org/10.3102/0034654315627366>
- Lee, D.K. 2016. Alternatives to P-value: confidence interval and effect size. *Korean J Anesthesiol*, 69(6): 555-562. <http://dx.doi.org/10.4097/kjae.2016.69.6.555>
- Lestari, I.D. 2019. Pengaruh model pembelajaran inkuiri terbimbing dikombinasikan dengan respon paper terhadap hasil belajar kognitif peserta didik kelas X SMAN 1

- Palas Lampung Selatan. *Undergraduate Thesis*, UIN Raden Intan Lampung, Indonesia.
- Lolombulan, J.H. 2017. *Statistika - bagi peneliti pendidikan*. ANDI, Yogyakarta.
- Mauritha, S., Nur, S., & Adlim. 2017. Pengaruh penggunaan model pembelajaran inkuiri terbimbing terhadap keterampilan proses sains siswa MAN. *Jurnal IPA dan Pembelajaran IPA*, 1(1): 1-5.
- Megawati, Y. 2018. Pengaruh model pembelajaran inkuiri terbimbing terhadap hasil belajar siswa pada mata pelajaran biologi kelas X di SMA Negeri 1 Semende Darat Ulu. *Undergraduate Thesis*, UIN Raden Fatah Palembang, Indonesia.
- Mertler, C.A. & Reinhart, R.V. 2017. *Advanced and multivariate statistical methods practical application and interpretation*, 6th edition, Routledge Taylor & Francis, New York.
- Mills, G.E. & Gay, L.R. 2019. *Educational research: competencies for analysis and application*, 12th edition, Pearson Education Limited, New York.
- Muslimah, S.M. 2019. Efektivitas sesi poster dengan metode gallery walk dalam pembelajaran inkuiri terbimbing terhadap hasil belajar biologi siswa pada konsep animalia. *Undergraduate Thesis*, UIN Syarif Hidayatullah Jakarta, Indonesia.
- Patten, M.L. & Newhart, M. 2017. *Understanding Research Methods: an Overview of the essentials*, 10th edition, Routledge, New York.
- Puspita, M.E., Setiadi, D., & Raksun, A. 2018. Perbedaan kemampuan berpikir kritis peserta didik dengan menggunakan model inkuiri terbimbing dan problem based learning (PBL) dalam pembelajaran biologi pada peserta didik kelas X SMAN 2 Gerung tahun ajaran 2016/2017. *Proceedings of The Seminar Nasional Pendidikan Biologi*, 2018 June, p.242-248.
- Richter, O.Z., Kerres, M., Bedenlier, S., & Bond, M. 2020. *Systematic reviews in educational research methodology, perspectives and application*, Buntins, K. (Eds.), Springer VS, Wiesbaden.
- Sadeh, I. & Zion, M. 2009. The development of dynamic inquiry performances within an open inquiry setting: a comparison to guided inquiry setting. *Journal of Research in Science Teaching*, 46(10):1137-1160. <http://dx.doi.org/10.1002/tea.20310>
- Sari, I.S., Lestari, S.R., & Sari, M.S. 2020. Development of a guided inquiry-based e-module on respiratory system content based on research results of the potential single garlic extract (*allium sativum*) to improve student creative thinking skills and cognitive learning outcome. *Jurnal Pendidikan Sains Indonesia*, 8(2):228-240. <http://dx.doi.org/10.24815/jpsi.v8i2.17065>
- Setiawati, E. 2018. Pengaruh model pembelajaran inkuiri terbimbing terhadap hasil belajar peserta didik pada konsep jamur. *Undergraduate Thesis*, UIN Syarif Hidayatullah Jakarta, Indonesia.
- Suwanto, S.A. 2017. Manajemen layanan repository perguruan tinggi. *Lentera Pustaka*, 3(2):165-176. <http://dx.doi.org/10.14710/lenpust.v3i2.16740>
- Tabachnick, B.G. & Fidell, L.S. 2019. *Using multivariate statistics*, 7th edition, Pearson, Boston.

- Tomczak, M. & Tomczak, E. 2014. The need to report effect size estimates revisited: an overview of some recommended measures of effect size. *TRENDS in Sport Sciences*, 1(21):19-25.
- UIN Syarif Hidayatullah Institutional Repository. 2020. Daftar skripsi Fakultas Ilmu Tarbiyah Dan Keguruan program studi Pendidikan Biologi 2010-2019. <http://repository.uinjkt.ac.id/dspace/handle/123456789/3>
- Ulfah, M., Hairida., Arifiyanti, F., Permasari, N., & Sabila, J.A. 2021. Analisis permasalahan pendidik ipa dalam proses penilaian pembelajaran. *Jurnal IPA dan Pembelajaran IPA*, 5(2):186-196. <http://dx.doi.org/10.24815/jipi.v5i2.21163>
- Wahono, R.S. 2015. A systematic literature review of software defect prediction: research trends, datasets, methods and frameworks. *Journal of Software Engineering*, 1(1):1-16.
- Yasniati. 2017. Penerapan model pembelajaran inkuiri terbimbing terhadap hasil belajar pada materi laju reaksi. *Jurnal Pendidikan Sains Indonesia*, 5(2):1-9. <http://dx.doi.org/10.24815/jpsi.v5i2.9807>
- Zarisa, A. & Saminan. 2017. Penerapan pembelajaran inkuiri menggunakan metode pictorial riddle pada materi alat-alat optik untuk meningkatkan kreativitas dan hasil belajar siswa. *Jurnal Pendidikan Sains Indonesia*, 5(1):1-4.
- Zulfiani, Z. & Herlanti, Y. 2018. Scientific inquiry perception and ability of pre-service teachers. *Journal of Turkish Science Education*, 15(1):128-140. <http://dx.doi.org/10.12973/tused.10225a>
- Zulfiani, Z., Yunistika, R., & Juanengsih, N. 2018. Enhancing students' higher-order thinking skills through guided and free inquiry-based learning. *Proceedings of the 3rd International Conference on Education in Muslim Society (ICEMS 2017)*, 115(6):28-31. <http://dx.doi.org/10.2991/icems-17.2018.6>