Jurnal Teknologi Pendidikan, Desember 2023, 25 (3), 569-583
DOI: http://dx.doi.org/10.21009/JTP2001.6
p-ISSN: 1411-2744 e-ISSN: 2620-3081
Accredited by Directorate General of Strengthening for Research and Development
Available online at http://journal.unj.ac.id/unj/index.php/jtp

Meta-Analysis of The Effectiveness of The Flipped Classroom Model On Students' HOTS in Mathematics

Pujiriyanto[1(*)], Calista Devi Handaru[2], Julham Hukom[3]
[1]Department of Educational Technology, Faculty of Education and Psychology, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia
[2]Department of Educational Science, Postgraduate School (SPs), Universitas Negeri Yogyakarta, Yogyakarta, Indonesia
[3]Department of Arabic Language Education, Faculty of Languages and Letters, Universitas Negeri Makassar, Makassar, Indonesia

Received:
Revised:
Accepted:

Abstract
The use of the Flipped Classroom (FC) Model has become a global phenomenon in the last decade, and has had a significant impact on mathematics learning objectives. Nevertheless, the scientific literature shows inconsistent results in testing the underlying theoretical assumptions in this model. This meta-analysis study collected analyses of 42 effect sizes from 31 journal articles and international proceedings between 2014 and 2024. The Comprehensive Meta-Analysis (CMA) program was used as a calculation tool, with the Hedges coefficient used to calculate the effect size at a 95% confidence level. The results of analysis using the random-effect estimation model show that the overall effect size is 1.06 (large effect category). This shows that the use of the FC model has a relatively high positive influence on High Order Thinking Skills (HOTS) mathematics when compared to the conventional model. In addition, analysis of research characteristics shows that the effectiveness of using the FC model can be increased by considering factors such as the type of platform used, the type of FC model applied, educational level, and the capacity of the experimental class. These findings provide valuable guidance for educators in the future use of the FC model, particularly in developing students' mathematics HOTS.

Keywords: Flipped Classroom; HOTS; Mathematics; Meta Analysis; Effect Size

(*) Corresponding Author: pujiriyanto@uny.ac.id

How to Cite: Pujiriyanto, Handaru, C., & Hukom, J. (2024). Meta-Analysis of The Effectiveness of The Flipped Classroom Model On Students' HOTS in Mathematics. *JTP - Jurnal Teknologi Pendidikan*, *25*(3), 569-583. https://doi.org/10.21009/jtp.v25i3.45848

INTRODUCTION

In the world of education, there are two very simple reasons why HOTS is important, firstly students must be successful in school and secondly students will grow into adults who make positive contributions to society (Conklin, 2012). HOTS can be trained on students in all disciplines, including mathematics. Research results confirm that HOTS has a strong correlation with academic success (Rahmawatiningrum et al., 2019; Siburian et al., 2019). In addition, higher-order thinking skills enable students to better deal with complex problems and unexpected situations, because they have been trained to analyze, evaluate, and create creative solutions (Azid et al., 2022).

In the last ten years, the Flipped Classroom (FC) model has become one of the learning strategies that has been adopted in education systems in various countries. With the advancement of digital technology, the FC model has transcended the boundaries of the traditional classroom and provided broader educational access (Herzog & Casale, 2022; Lepper & Malone, 2021). The FC model is often associated with improving academic mathematics abilities because teachers can provide opportunities for students to study at home at their own pace (Anugrah et al., 2020; Jesionkowska et al., 2020; Lo et al., 2017; Maryati et al. , 2022; Susilawati & Khaira, 2021; Wei et al., 2020), encouraging students to take responsibility for their own learning (Demitriadou et al., 2020), enabling better personalization of learning (Panjaburee et al., 2022) , increasing student interest and motivation (Cobena & Surjono, 2022). The development of the FC model also provides students with the opportunity to practice problem solving skills and find alternative solutions. This is due to flexible learning which provides broad access to various learning resources (Kaput et al., 2020; Lin, 2019).

A number of studies have been conducted to evaluate the effectiveness of the FC model against the mathematical HOTS of students at different levels of education. However, the results were inconsistent. Some studies show that the use of the FC model is more effective in improving students' mathematical HOTS when compared to conventional learning models for example (Albalawi, 2018; Casem, 2016; Hanifah et al., 2023; Jarrah & Diab, 2019; Lo & Hew, 2020; Zebidi, 2021; Zeineddine, 2018). However, other studies have found that using the FC model does not provide better results (Esperanza et al., 2016; Ramadhani et al., 2019). The difference in the results of this study will confuse education policy makers, especially mathematics teachers who require accurate and consistent information, making it difficult to determine the right conditions for using the FC model to achieve a more optimal level of effectiveness.

The problem of inconsistency of the above research results can be overcome by collecting and summarizing various primary research results. Quantitative research methods such as meta-analysis studies can be used to provide more accurate information in policy making (Muhtadi et al., 2022; Kamsurya et al., 2022; Martaputri et al., 2021; Setiawan et al., 2022; Samritin et al., 2023). Meta-analysis studies have a special role in integrating findings from various major studies and identifying reasons for variation in outcomes that need to be considered in future practice (Mullen, 2001; Wicherts, 2020). Compared to a single experimental study, meta-analyses can provide in-depth and accurate conclusions about the two variables investigated (Borenstein et al., 2021). Therefore, meta-analysis studies are needed to draw more conclusive conclusions about the applicability of the FC model to students' mathematical HOTS, so that teachers can consider its future application. Meta-analysis research related to the effect of using the FC model has previously been conducted by (Güler et al., 2023; Purnomo et al., 2022; Sulistyowati et al., 2023).

However, their analysis was limited to the general impact of FC use on mathematical achievement, without investigating specifically its effect on Higher HOTS. Previous meta-analyses have also been conducted by (Wagner et al., 2021; Yakar, 2021), but their focus has been limited to the effects of FC use at primary and secondary education levels. Last but not least, important characteristics of the study, such as the type of platform used, the capacity of the experimental class, and the variety of FC models applied, have not been incorporated into their analysis. These limitations markedly reduce our ability to understand holistically the effectiveness of using the FC model in the context of overall mathematics learning. Thus, there is an urgent need for more comprehensive research that considers not only mathematical achievement in general, but also critical aspects such as HOTS ability, as well as taking into account the characteristics of more detailed studies to generate a deeper understanding of the role of the FC model in mathematics learning.

This study intends to complement the limitations of previous research with the aim of (1) assessing the effectiveness of using the FC model on students' overall mathematical HOTS and (2) knowing the difference in the effectiveness of the FC model based on research characteristics. This is necessary to help educators consider the most optimal conditions for using the FC model in improving students' mathematical HOTS.

METHODS

Research Design and Procedures

This research uses a meta-analysis design to test the effect of the FC model on students' mathematics skills. This design was chosen to integrate findings that have been produced by previous studies, thereby enabling a holistic and general picture of the impact of the FC model to be presented.

Research Question

This goal will be achieved by analysing the primary studies conducted on the impact of FC models on students' mathematical HOTS using meta-analysis studies. The following research questions are in focus:

1) Does the use of FC models produce a significant effect on mathematical HOTS compared to conventional models?
2) Are there differences in the effectiveness of using the FC model based on the following platform types: (a) Edmodo, (b) Google Classroom, (c) Khan Academy, (d) Moodle, and (e) WhatsApp Group?
3) Are there differences in the effectiveness of using FC models based on the following platform types: (a) conventional FC, (b) integrated FC and cultural learning, (c) integrated FC and cultural learning, (d) integrated FC and GeoGebra, and (e) integrated FC and problem based learning?
4) Are there differences in the effectiveness of using the FC model based on the following education levels: (a) Elementary School, (b) Junior High School, (c) Senior High School, and (d) College?
5) Are there differences in the effectiveness of using the FC model based on the capacity of the following experimental classes: (a) less or equal to 30, and (b) more than 30?

Inclusion Criteria and data collection

The inclusion criteria are used in meta-analysis as a filter to collect research artifacts that are of good quality and have adequate content completeness. Data obtained from artifacts can be scientifically justified. The inclusion criteria for this research include studies that meet the following requirements:

1) Research published between 2014 and 2024;
2) Students at different educational levels participating in mathematics learning using the FC model;
3) Experimental, quasi-experimental, or observational research with a clearly defined control group;
4) Studies reporting data on student HOTS resulting from intervention using the FC model;
5) Studies available in English to facilitate comprehension and analysis; and
6) Studies must provide adequate statistical data for calculating effect sizes.

Data collection refers to the specified inclusion criteria. Literature collection uses the Google Scholar database and the Google search engine. Keywords used included "Flipped Classroom" AND "Mathematics High Order Thinking Skills". In addition, literature selection was carried out to obtain information from related sources that may not be included in the database. The screening results obtained from 31 primary studies were used as material for meta-analysis. The screening process can be seen in Figure 1 below.

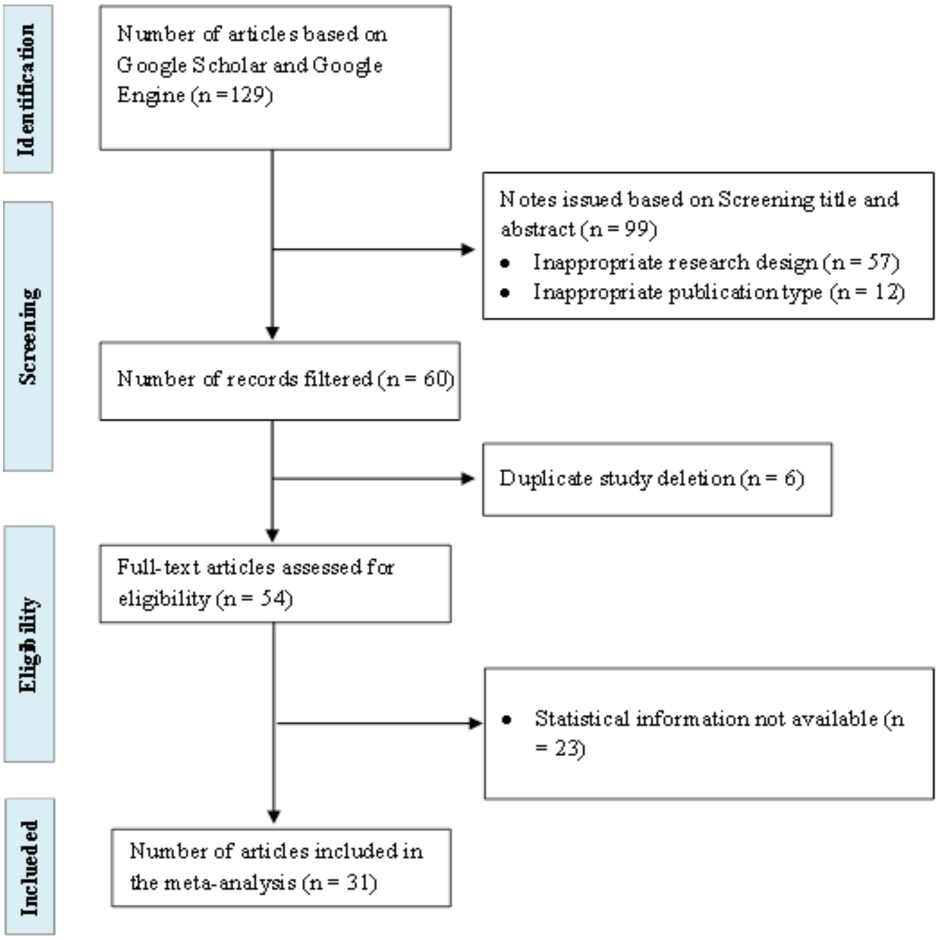

Figure 1. Results of the literature screening process

There were several studies involving more than one control group, resulting in 42 effect sizes being analyzed. More details can be seen in Table 1 below.

Table 1. Coding of Studies that Meet the Inclusion Criteria

Coding Content	Moderator Variable	Frequency
Platforms Type	Moodle	1
	Edmodo	1
	Google Classroom	7
	Khan Academy	4
	WhatsApp	15
	Not reported	8
Model FC Type	Conventional FC	37
	Integrated FC and Cltural Learning (FC-CL)	1
	Integrated FC and Game based learning (FC-GbL)	1
	Integrated FC and GeoGebra (FC-G)	2
	Integrated FC and PBL (FC-PBL)	1
Educational Level	Elementary School (ES)	2
	Junior High School (JHS)	13
	Senior High School (SHS)	19
	College	8
Class Capacity	Small (≤ 30)	14
	Large (> 30)	28

Data Analysis

Data analysis used CMA version 3 software. The meta-analysis scheme applied involved several stages as shown in Figure 1 below.

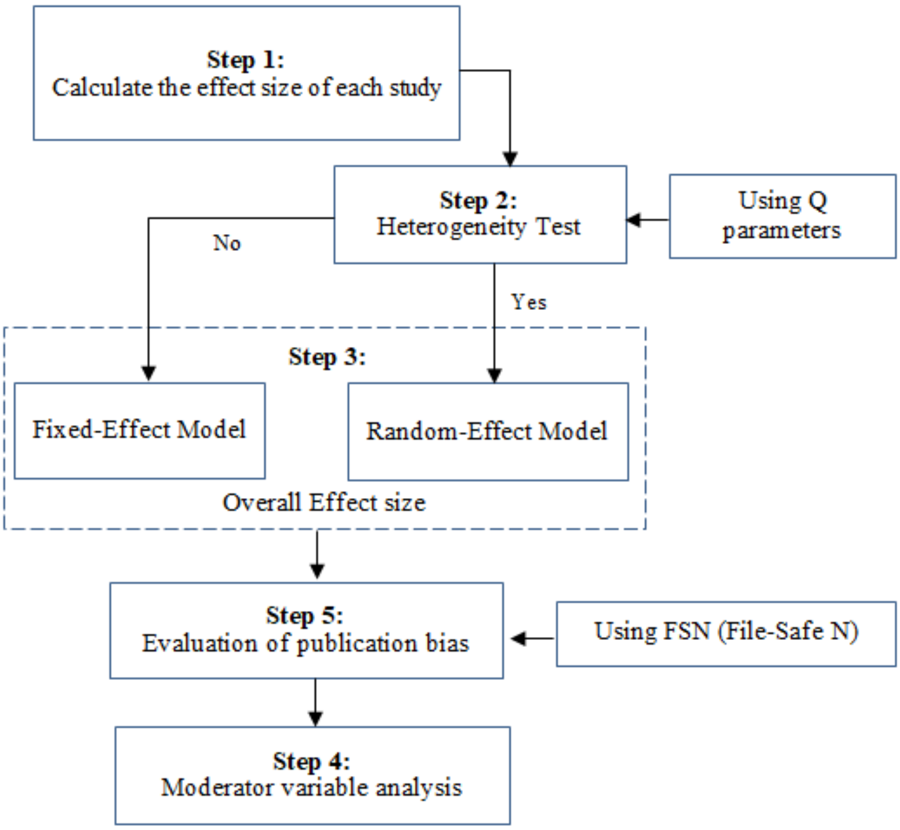

Figure 1. Data analysis scheme using CMA

The process of interpreting effect sizes in this research refers to the classification proposed by (Cohen et al., 2002). The effect size (g) categories based on this classification are: 1) The "ignored" category if the effect size value is $0.00 < g \leq 0.19$; 2) Category "small effect" if the effect size value is $0.19 < g \leq 0.49$; 3) Category "medium effect" if the effect size value is $0.49 < g \leq 0.79$; 4) Category "large effect" if the effect size value is $0.79 < g \leq 1.29$; and 5) Category "very large effect" if the effect size value is $g > 1.29$.

RESULTS

First Research Question

The first objective of this analysis was to assess the effect of FC use on students' overall math HOTS. Therefore, based on CMA-assisted calculations, effect sizes and confidence interval limits are obtained and presented in Figure 2.

Study name	Statistics for each study						
	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Albawi (2018)	1,965	0,253	0,064	1,471	2,460	7,783	0,000
Andriani et al. (2022) a	1,513	0,353	0,125	0,821	2,205	4,284	0,000
Andriani et al. (2022) b	0,526	0,315	0,100	-0,092	1,144	1,667	0,095
Ario & Asra (2018)	1,714	0,506	0,256	0,722	2,705	3,388	0,001
Arnawa & Setiawan (2021)	0,474	0,147	0,021	0,187	0,761	3,235	0,001
Casem (2016)	0,776	0,410	0,168	-0,027	1,579	1,894	0,058
Esperanza et al (2016)	0,424	0,210	0,044	0,012	0,836	2,015	0,044
Hanifah et al. (2023) a	1,898	0,290	0,084	1,330	2,466	6,549	0,000
Hanifah et al. (2023) b	0,905	0,252	0,064	0,411	1,399	3,590	0,000
Jarah & Diab (2019) a	0,977	0,236	0,056	0,514	1,439	4,139	0,000
Jarah & Diab (2019) b	0,971	0,236	0,056	0,509	1,433	4,117	0,000
Jarah & Diab (2019) c	0,908	0,234	0,055	0,449	1,367	3,876	0,000
Jarah & Diab (2019) d	0,192	0,223	0,050	-0,246	0,630	0,860	0,390
Juniantari et al (2018)	1,198	0,265	0,070	0,679	1,717	4,526	0,000
Khofifah et al (2021) 1a	1,435	0,282	0,080	0,883	1,988	5,090	0,000
Khofifah et al (2021) 1b	0,658	0,258	0,066	0,153	1,163	2,555	0,011
Khofifah et al (2021) 2a	1,409	0,281	0,079	0,858	1,959	5,015	0,000
Khofifah et al (2021) 2b	0,982	0,266	0,071	0,461	1,503	3,695	0,000
Khoirotunnisa & Irhadtanto (2020)	1,600	0,285	0,081	1,042	2,158	5,623	0,000
Lo & Hew (2018)	0,720	0,275	0,075	0,181	1,258	2,621	0,009
Makinde (2020)	0,920	0,127	0,016	0,672	1,168	7,257	0,000
Nida et al. (2019)	1,271	0,168	0,028	0,942	1,600	7,579	0,000
Nurfadillah (2022)	0,993	0,280	0,078	0,445	1,542	3,551	0,000
Pinontoan & Walean (2019)	0,996	0,322	0,103	0,366	1,627	3,097	0,002
Pratiwi (2021)	1,790	0,293	0,086	1,215	2,365	6,104	0,000
Ramadhani (2019)	0,108	0,252	0,063	-0,385	0,601	0,430	0,667
Sihotang et al. (2023)	0,816	0,309	0,095	0,211	1,421	2,644	0,008
Spotts & Blumme (2020) a	1,106	0,319	0,102	0,481	1,730	3,469	0,001
Spotts & Blumme (2020) b	0,697	0,305	0,093	0,099	1,295	2,284	0,022
Syahrul et al (2020)	0,769	0,260	0,068	0,260	1,279	2,958	0,003
Wei et al (2020)	0,621	0,216	0,047	0,197	1,046	2,870	0,004
Zatalini et (2017)	0,630	0,266	0,071	0,109	1,150	2,369	0,018
Zebidi (2021)	3,466	0,445	0,198	2,594	4,337	7,796	0,000
Zineddine (2018)	0,424	0,311	0,097	-0,185	1,033	1,366	0,172
Salsabila & Maarif (2022)	0,730	0,263	0,069	0,213	1,246	2,770	0,006
Darwani et al. (2023)	1,200	0,316	0,100	0,581	1,819	3,801	0,000
Kiptiyah et al (2021) b	0,788	0,217	0,047	0,362	1,214	3,629	0,000
Pratidiana et al. (2022)	1,519	0,353	0,125	0,826	2,211	4,297	0,000
Saputra & Mujib (2018)	3,117	0,362	0,131	2,407	3,827	8,607	0,000
Mubarokah et al. (2022)	0,742	0,247	0,061	0,258	1,226	3,006	0,003
Arifin & Herman (2018)	1,003	0,275	0,076	0,463	1,542	3,642	0,000
Nurhayati (2022)	0,891	0,287	0,082	0,329	1,453	3,106	0,002

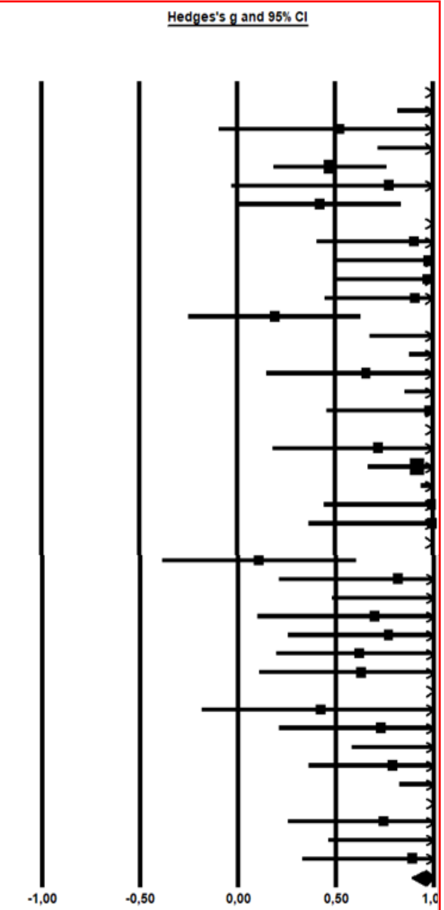

Figure 2. Forest Plot

Figure 1 shows the overall effect size ranging between 0.11 and 3.47, with a 95% confidence limit, while Figure 3 shows the level of effect size of the entire study based on Cohen et al. (2002).

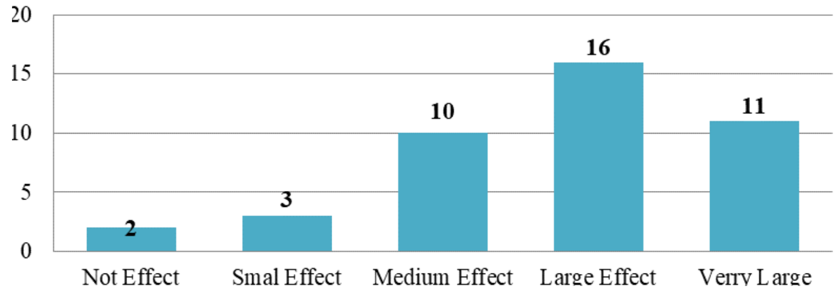

Figure 3. Classification of Effect Sizes

Figure 3 shows the difference in effect size obtained from research conducted on the use of FC on mathematics HOTS, while Table 2 illustrates the results of descriptive meta-analysis according to the estimation method.

Tabel 2. Example Table

Estimation Model	N	g	P	Df	Heterogeneity Q	p	Decision
Random-Effect	42	1.06	< 0.01	41	180.210	0.00	Reject H_0
Fixed-Effect	42	0.97	< 0.01	41			

The estimation method was determined through an effect size homogeneity test which showed a Q value of 180.210 and found more than 74.56 (df = 41; p = 0.05) in table χ2. That is, the size of the effect between studies is different. Therefore, an estimation model to determine the impact of FC use on students' overall mathematical HOTS can be evaluated using a random-effects model. The random-effects model in Table 3 shows the average effect size is 1.06 at a 95% confidence interval. This effect size is quite large. In addition, the significance test results give a value of p = 0.00 which shows that the use of FC against mathematical HOTS produces a larger effect size than conventional approaches.

Next, an evaluation of publication bias is investigated to identify whether there is systematic bias in the selection of studies to be included in the meta-analysis. This can occur when studies with significant or interesting results are more likely to be published than studies with insignificant or less interesting results. This evaluation helps avoid drawing inaccurate or biased conclusions. Therefore, the Fail-Safe N (FSN) approach was included in determining the presence of publication bias in this study, and the results are presented in Table 3.

Tabel 3. FSN Results

P	0.00
Alpha	0.05
z for Alpha	1.96
N	42
P > number of missing studies	6888

With target significance values ($\alpha = 0.05$) and $p < 0.001$, these results show that despite a number of unpublished studies, the statistically significant combined effect is maintained. Therefore, the conclusion that can be drawn is that this meta-analysis has a high level of security against the problem of publication bias, and the results can be considered scientifically valid. In other words, even if there are additional studies that have not yet been published or the results are not significant, this will not change the conclusions of this meta-analysis as a whole.

Research Results Regarding the Second Question

Descriptive statistics about the second question are illustrated in Table 4.

Tabel 4. Effect Size by Platform Type

Platform Type	N	Effect Size	Std. Error	P	Heterogeneity df	Qb	P	Decision
Moodle	6	1.06	0.11	0.00	5	21.99	0.00	Reject H_0
Google Classroom	7	0.70	0.09	0.00				
Edmodo	1	1.20	0.26	0.00				
Khan Academy	4	0.75	0.12	0.00				
WhatsApp	15	0.74	0.06	0.00				
Not Report	8	1.04	0.11	0.00				

Table 4 obtained the statistical value of Q homogeneity test is 21.99, this value is greater than 12.59 (df = 5; p = 0.05) in table $\chi 2$. These results indicate that the effectiveness of the FC model on HOTS mathematics differs significantly based on type of platform.

Research Results Regarding the Third Question

Descriptive statistics about the third question are illustrated in Table 5.

Tabel 5. Effect Size by FC Model Type

Model FC Type	N	Effect Size	Std. Error	P	Heterogeneity df	Qb	P	Decision
Conventional FC	37	0.79	0.04	0.00	4	25.12	0.00	Reject H_0
FC-CL	1	0.82	0.22	0.00				
FC-GbL	1	0.92	0.27	0.01				
FC-G	2	1.74	0.22	0.00				
FC-PBL	1	0.81	0.25	0.00				

Table 5 obtained the statistical value of Q homogeneity test is 25.12, this value is greater than 9.49 (df = 4; p = 0.05) in table $\chi 2$. These results indicate that the effectiveness of the FC model on HOTS mathematics differs significantly based on type of FC model.

Research Results on the Fourth Question

Descriptive statistics about the fourth question are illustrated in Table 6.

Tabel 6. Effect Size Based on Education Level

Level of Education	N	Effect Size	Std. Error	P	Heterogeneity df	Qb	P	Decision
ES	2	0.86	0.19	0.00	3	20.46	0.00	Reject H_0
SHS	19	0.87	0.06	0.00				
JHS	13	1.11	0.07	0.00				
College	8	0.99	0.09	0.00				

Table 6 obtained the statistical value of Q homogeneity test is 20.46, this value is greater than 7.82 (df = 3; p = 0.05) in table $\chi 2$. These results indicate that the effectiveness of the FC model on HOTS mathematics differs significantly based on the level of education.

Research Results on Question Fifth

Descriptive statistics about the fifth question are illustrated in Table 7.

Tabel 7. Effect Size by Experimental Class Capacity Type

Class Capacity	N	Effect Size	Std. Error	P	Heterogeneity df	Heterogeneity Qb	Heterogeneity P	Decision
≤ 30	14	1.38	0.03	< 0.01	1	20.29	0.00	Reject H_0
> 30	28	0.87	0.07	< 0.01				

Table 7 obtained statistical value Q homogeneity test is 20.46, this value is greater than 7.82 (df = 3; p = 0.05) in table $\chi 2$. These results indicate that the effectiveness of the FC model on HOTS mathematics differs significantly based on class capacity.

DISCUSSION

This meta-analysis study analyzed 42 effect measures. The combined effect size was 1.06. This value indicates that the use of FC has a greater and significant impact compared to conventional models in improving students' math HOTS. It also indicates that on average students who get FC treatment have math HOTS results about 84% higher than students who are in conventional classes, who initially have an equivalent level of ability. Based on the mentioned interpretation table from (Coe, 2002), the average student who ranked 15th in the FC experimental group was equivalent to the student who was ranked 5th in the control group. This suggests that the use of FC has a significant impact in improving HOTS mathematics. The findings of this meta-analysis are consistent with findings from several previous meta-analyses conducted by (Cheng et al., 2019; Purnomo et al., 2022; Wagner et al., 2021; Yakar, 2021), which also found that the application of the FC model contributes significantly to student learning outcomes.

The consistency of these research findings provides stronger strength and credibility in recommending the application of the FC model as an effective approach to improving students' mathematics HOTS. The results of the analysis of research characteristics show that the FC model used will be more effective by considering the type of platform, type of FC model, level of education, and capacity of the experimental class. Further explanation of each research characteristic studied is explained in the next paragraph.

First, the results of our analysis show that the platform type variable moderates the impact of using the FC model on mathematical HOTS. In other words, the effectiveness of using the FC model on student mathematics HOTS differs based on the type of platform used. The FC model proved to be the most effective in using the Moodle platform, followed by Edmodo, Khan Academy, WhatsApp, and Google Classroom. This finding is in line with the meta-analysis conducted by Sulistyowati et al. (2023), who found that the platform approach to using the flipped classroom model can provide different mathematics learning outcomes. Therefore, it is important for mathematics educators to pay attention to the right approach when practicing the flipped classroom model. Although there are significant differences between the types of platforms used, the results of the analysis unequivocally reveal that all types of platforms analyzed are effective for increasing students' mathematics HOTS.

Second, the results of our analysis also show that the FC model type variable moderates the impact of using the FC model on mathematics HOTS. In other words, the effectiveness of using the FC model on student mathematics HOTS differs based on the type of FC model used. The FC model was found to be most effective in the type of FC model integrated with GeoGebra (FC-G), followed by FC integration with game-based learning (FC-GbL), FC integration with cultural learning (FC-CL), FC integration with PBL (FC -PBL), and finally conventional FC. The meta-analysis by Purnomo et al. (2022) and Sulistyowati et al. (2023) aligns with this finding, indicating that varying learning approaches in the flipped classroom model can lead to varying mathematics performance. Therefore, it is important for mathematics educators to pay attention to appropriate learning approaches when practicing the flipped classroom model. Although there are significant differences between the types of FC models used, the analysis results show that all types of models have a significant effect on students' mathematics HOTS.

Third, the results of the analysis also show that the educational level variable moderates the influence of the FC model on students' mathematics HOTS. In other words, the effectiveness of using the FC model on students' mathematics HOTS differs based on certain educational levels. The results of the analysis show that the use of the FC model for student mathematics HOTS is most effective at higher education levels such as college and SHS compared to ES and SMP. This finding is in line with the results of a previous meta-analysis study conducted by Cheng et al. (2019), Güler et al. (2023), and Vitta and Al-Hoorie (2023) which show that the educational level variable moderates the magnitude of the influence of using the FC model on students' academic abilities. Even though there are significant differences between educational level groups, the use of the FC model has a significant influence on students' mathematics HOTS at all levels of education.

Finally, the results of our meta-analysis show that the classroom capacity variable moderates the effect of using the FC model on students' mathematics HOTS. In other words, the effectiveness of using the FC model for student mathematics HOTS differs based on class capacity. This can be seen from the experimental class capacity ≤ 30 having a larger average effect size compared to the experimental class capacity > 30. This shows that a smaller experimental class capacity will likely produce a more optimal impact in using the FC model. This finding is in line with several previous studies conducted (Mawardi et al., 2023; Samritin et al., 2023; Turgut & Turgut, 2018), which stated that small study groups tend to provide greater effectiveness. Even though there are significant differences between class capacity groups, the results of the analysis show that the use of the FC model is effective in each group analyzed.

CONCLUSION

This research was conducted by integrating the findings of the influence of using the FC model on students' mathematics HOTS, both overall and on several research moderator variables. The results of the analysis reveal that the use of the FC model has a positive impact on students' mathematics HOTS. Assessment of FC effectiveness based on moderator variables shows that the effectiveness of using the FC model for HOTS mathematics differs based on platform type, FC model type, educational level, and class capacity.

Although the use of the FC model was confirmed to produce a great effect on students' mathematical HOTS, the results were based only on studies with certain criteria, and some similar studies were not analysed because the required statistical information was inadequate. For the purpose of this study, only four characteristics of the study were examined which included platform type, FC model type, education level, and class capacity. While some others such as learning materials and the length of treatment have not been included for analysis. As a result, these conclusions do not reflect the effectiveness of using the FC model in overall mathematics learning. Therefore, further research requires an in-depth investigation to determine the effectiveness of the FC model using several characteristics that have not been studied.

REFERENCES

Albalawi, A. S. (2018). The Effect of Using Flipped Classroom in Teaching Calculus on Students' Achievements at University of Tabuk. *International Journal of Research in Education and Science*, 198–207. https://doi.org/10.21890/ijres.383137

Anugrah, A., Ibrahim, N., & Sukardjo, M. (2020). How Flipped Classroom Helps the Learning in the Times of Covid-19 Era?. *JTP-Jurnal Teknologi Pendidikan*, *22*(3), 151-158.

Arlinwibowo, J., Retnawati, H., & Kartowagiran, B. (2022). The impact of ICT utilization to improve the learning outcome: A meta-analysis. *International Journal of Evaluation and Research in Education (IJERE)*, *11*(2), 522–531.

Azid, N., Ali, R. M., El Khuluqo, I., Purwanto, S. E., & Susanti, E. N. (2022). Higher Order Thinking Skills, School-Based Assessment and Students' Mathematics Achievement: Understanding Teachers' Thoughts. *International Journal of Evaluation and Research in Education*, *11*(1), 290-302.

Borenstein, M., Hedges, L. V, Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. John Wiley & Sons.

Casem, R. Q. (2016). Effects of flipped instruction on the performance and attitude of high school students in mathematics. *European Journal of STEM Education*, *1*(2), 37–44. https://doi.org/10.20897/lectito.201620

Cheng, L., Ritzhaupt, A. D., & Antonenko, P. (2019). Effects of the flipped classroom instructional strategy on students' learning outcomes: a meta-analysis. *Educational Technology Research and Development*, *67*(4), 793–824. https://doi.org/10.1007/s11423-018-9633-7

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review*, *9*, 88–113. https://doi.org/10.1016/j.edurev.2013.01.001

Cobena, D. Y., & Surjono, H. D. (2022). Implementation of Flipped Classroom Model in Vocational High School: A Systematic Literature Review. *JTP - Jurnal Teknologi Pendidikan*, *24*(1), 79-92.

Coe, R. (2002). It's the effect size, stupid: What effect size is and why its important. *Annual Conference of the British Educational Research Association, Exeter: Devon. Education-Line*, *18*.

Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education*. routledge.

Conklin W. (2012). Higher order thinking skills to develop 21st century learners. California: Shell Education Publishing.

Demitriadou, E., Stavroulia, K.-E., & Lanitis, A. (2020). Comparative evaluation of virtual and augmented reality for teaching mathematics in primary education. *Education and Information Technologies*, *25*(1), 381–401.

Esperanza, P., Fabian, K., & Toto, C. (2016). *Flipped Classroom Model: Effects on Performance, Attitudes and Perceptions in High School Algebra* (pp. 85–97). https://doi.org/10.1007/978-3-319-45153-4_7

Güler, M., Kokoç, M., & Önder Bütüner, S. (2023). Does a flipped classroom model work in mathematics education? A meta-analysis. *Education and Information Technologies*, *28*(1), 57–79. https://doi.org/10.1007/s10639-022-11143-z

Hanifah, M., Ishartono, N., Halili, S. H. B., & Razak, R. B. A. (2023). *Improving secondary school student's mathematics critical thinking in geometry transformation through GeoGebra-Based flipped learning: An experimental study*. 020024. https://doi.org/10.1063/5.0141401

Herzog, M., & Casale, G. (2022). The Effects of a Computer-Based Mathematics Intervention in Primary School Students with and Without Emotional and Behavioral Difficulties. *International Electronic Journal of Elementary Education*, *14*(3), 303–317. https://doi.org/10.26822/iejee.2022.246

Jarrah, A. M., & Diab, K. M. A. B. M. (2019). The Effect of Flipped Classroom Model on Students' Achievement in the New 2016 Scholastic Assessment Test Mathematics Skills. *The Journal of Social Sciences Research*, *53*, 769–777. https://doi.org/10.32861/jssr.53.769.777

Jesionkowska, J., Wild, F., & Deval, Y. (2020). Active Learning Augmented Reality for STEAM Education—A Case Study. *Education Sciences*, *10*(8), 198. https://doi.org/10.3390/educsci10080198

Kamsurya, M. A., Wijaya, A., Ramadhani, R., & Hukom, J. (2022). The Effect of Self-Efficacy on Students' Mathematical Abilities: A Meta-Analysis Study. *Jurnal Pendidikan Progresif,* 12(2), 451-463.

Kaput, J., Hegedus, S., & Lesh, R. (2020). Technology becoming infrastructural in mathematics education. In *Foundations for the future in mathematics education* (pp. 173–191). Routledge.

Lepper, M. R., & Malone, T. W. (2021). Intrinsic motivation and instructional effectiveness in computer-based education. In *Aptitude, learning, and instruction* (pp. 255–286). Routledge.

Lin, Y.-T. (2019). Impacts of a flipped classroom with a smart learning diagnosis system on students' learning performance, perception, and problem solving ability in a software engineering course. *Computers in Human Behavior*, *95*, 187–196. https://doi.org/10.1016/j.chb.2018.11.036

Lo, C. K., & Hew, K. F. (2020). A comparison of flipped learning with gamification, traditional learning, and online independent study: the effects on students' mathematics achievement and cognitive engagement. *Interactive Learning Environments*, *28*(4), 464–481. https://doi.org/10.1080/10494820.2018.1541910

Lo, C. K., Hew, K. F., & Chen, G. (2017). Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education. *Educational Research Review*, *22*, 50–73. https://doi.org/10.1016/j.edurev.2017.08.002

Martaputri, N. A., Muhtadi, A., Hukom, J., & Samal, D. (2021). The Correlation between Emotional Intelligence and Academic Achievement: A Meta Analysis Study. *Jurnal Pendidikan Progresif*, 11(3), 511-523.

Maryati, L., Rafli, Z., & Setiadi, S. (2022). Effect of Blended Learning in high school Japanese Language Learning outcomes. *JTP-Jurnal Teknologi Pendidikan*, *24*(1), 105-119.

Mawardi, D. N., Budiningsih, C. A., & Sugiman. (2023). Blended Learning Effect on Mathematical Skills: A Meta-Analysis Study. *Ingénierie Des Systèmes d Information*, *28*(1), 197–204. https://doi.org/10.18280/isi.280122

Muhtadi, A., Pujiriyanto., Syafruddin, K., Hukom, J., & Samal, D. (2022). A meta-analysis: Emotional intelligence and its effect on mathematics achievement. *International Journal of Instruction*, 15(4), 745-762.

Muhtadi, A., Assagaf, G & Hukom, J. (2022). Self-efficacy and students' mathematics learning ability in Indonesia: A meta analysis study. International *Journal of Instruction*, 15(3), 1131-1146.

Mullen, B., Muellerleile, P., & Bryant, B. (2001). Cumulative metaanalysis: A consideration of indicators of sufficiency and stability. Personality and Social Psychology Bulletin, 27(11), 1450–1462.

Mulyono, D., & Hidayati, A. N. (2020). Peningkatan Hasil Belajar Mata Kuliah Media Pembelajaran Matematika Melalui Flipped Clasroom Berbantuan Schoology. *JTP-Jurnal Teknologi Pendidikan*, *22*(2), 88-95.

Panjaburee, P., Komalawardhana, N., & Ingkavara, T. (2022). Acceptance of personalized e-learning systems: a case study of concept-effect relationship approach on science, technology, and mathematics courses. *Journal of Computers in Education*, *9*(4), 681–705.

Purnomo, B., Muhtadi, A., Ramadhani, R., Manaf, A., & Hukom, J. (2022). The Effect of Flipped Classroom Model on Mathematical Ability: A Meta Analysis Study. *Jurnal Pendidikan Progresif*, *12*(3), 1201–1217. https://doi.org/10.23960/jpp.v12.i3.202216

Rahmawatiningrum, A., Kusmayadi, T. A., & Fitriana, L. (2019, October). Student's ability in solving higher order thinking skills (HOTS) mathematics problem based on learning achievement. In *Journal of Physics: Conference Series* (Vol. 1318, No. 1, p. 012090). IOP Publishing.

Ramadhani, R., Umam, R., Abdurrahman, A., & Syazali, M. (2019). The Effect of Flipped-Problem Based Learning Model Integrated With LMS-Google Classroom for Senior High School Students. *Journal for the Education of Gifted Young Scientists*, *7*(2), 137–158. https://doi.org/10.17478/jegys.548350

Samritin, S., Susanto, A., Manaf, A., & Hukom, J. (2023). A meta-analysis study of the effect of the blended learning model on students' mathematics learning achievement. *Jurnal Elemen*, *9*(1), 15–30. https://doi.org/10.29408/jel.v9i1.6141

Setiawan, A. A., Muhtadi, A., & Hukom, J. (2022). Blended Learning and Student Mathematics Ability in Indonesia: A Meta-Analysis Study. *International Journal of Instruction*, *15*(2), 905–916. https://doi.org/10.29333/iji.2022.15249a

Sıburıan, J., Corebıma, A. D., & Saptasarı, M. (2019). The correlation between critical and creative thinking skills on cognitive learning results. *Eurasian Journal of Educational Research*, *19*(81), 99-114.

Sulistyowati, E., Hukom, J., & Muhtadi, A. (2023). Meta-Analysis of Flipped Classroom on Students' Mathematics Abilities: Effectiveness and Heterogeneity Analysis. *JTP - Jurnal Teknologi Pendidikan*, *25*(2), 140–159. https://doi.org/10.21009/jtp.v25i2.27433

Susilawati, E., & Khaira, I. (2021). Higher order thinking skill (hots) dan model pembelajaran tpack serta penerapannya pada matakuliah strategi pembelajaran ppkn. *Jurnal Teknologi Pendidikan (JTP)*, *14*(2), 139-147.

Turgut, İ. G., & Turgut, S. (2018). The Effects of Visualization on Mathematics Achievement in Reference to Thesis Studies Conducted in Turkey: A Meta-Analysis. *Universal Journal of Educational Research*, *6*(5), 1094–1106. https://doi.org/10.13189/ujer.2018.060531

Vitta, J. P., & Al-Hoorie, A. H. (2023). The flipped classroom in second language learning: A meta-analysis. *Language Teaching Research*, *27*(5), 1268–1292. https://doi.org/10.1177/1362168820981403

Wagner, M., Gegenfurtner, A., & Urhahne, D. (2021). Effectiveness of the Flipped Classroom on Student Achievement in Secondary Education: A Meta-Analysis. *Zeitschrift Für Pädagogische Psychologie*, *35*(1), 11–31. https://doi.org/10.1024/1010-0652/a000274

Wei, X., Cheng, I.-L., Chen, N.-S., Yang, X., Liu, Y., Dong, Y., Zhai, X., & Kinshuk. (2020). Effect of the flipped classroom on the mathematics performance of middle school students. *Educational Technology Research and Development*, *68*(3), 1461–1484. https://doi.org/10.1007/s11423-020-09752-x

Wicherts, J. (2020). Meta-analysis. In *Encyclopedia of Personality and Individual Differences* (pp. 2859–2860). Springer International Publishing. https://doi.org/10.1007/978-3-319-24612-3_1326

Yakar, Z. Y. (2021). The Effect of Flipped Learning Model on Primary and Secondary School Students' Mathematics Achievement: A Meta-Analysis

Study. *Cukurova University Faculty of Education Journal*, *50*(2), 1329–1366. https://doi.org/10.14812/cuefd.865337

Zebidi, A. (2021). The impact of Flipped Classroom Strategy of Teaching Mathematics on Students' Achievements at Umm Al-Qura University. *Journal of Education*, *2*(81), 49-67. https://doi.org/10.21608/edusohag.2021.130232

Zeineddine, D. (2018). Investigate the Effects of Flipped Learning on Understanding of Mathematics for Secondary Students. *Journal of Mathematics Education*, *11*(1), 62–80. https://doi.org/10.26711/007577152790021