Kasetsart Journal of Social Sciences 44 (2023) 517–526

Kasetsart Journal of Social Sciences

journal homepage: http://kjss.kasetsart.org

Teachers as co-designers in education: A meta-analysis of the influence of co-designed teaching and learning on student outcomes

Watcharasak Sudla*, Chayut Piromsombat†

Department of Educational Research and Psychology, Faculty of Education, Chulalongkorn University, Bangkok 10330, Thailand

Article Info

Article history:
Received 9 March 2022
Revised 26 August 2022
Accepted 30 August 2022
Available online 21 June 2023

Keywords:
co-design,
co-design methodology,
meta-analysis,
student outcome

Abstract

Co-design is a potentially effective technique for supporting stakeholders in the development of educational innovations. Even though this approach is becoming increasingly popular in the education sector in Thailand, evidence to illustrate the current state of the co-design process is required. The purpose of this research is to analyze and synthesize educational primary studies that have used the co-design process in relation to educational innovations. Eight experimental and quasi-experimental studies were used to synthesize and examine the influence on student outcomes of conducting learning activities involving designed interventions. The results of the meta-analysis revealed that the learning activities implemented by designed interventions across the studies appeared to have heterogeneous effect sizes. These effects varied depending on the characteristics of the research design. Based on the findings of this study, some recommendations for using the co-design process to improve student outcomes in the Thai educational context are suggested.

Introduction

Co-design is the process of involving stakeholders of products and services in the design process, with the purpose of resulting in improvements and innovation (Burkett, 2019). Its features enable stakeholders in brainstorming to develop potential methods for solving problems and creating better solutions (Zamenopoulos & Alexiou, 2018). Co-design, also known as participatory design, was established in Scandinavia prior to the use of this term in various contexts (Sanders & Stappers, 2008). Its idea has been often used in business, particularly as a marketing activity, because designs or services that emphasize the experience of diverse customers can help to generate more value for goods and services.

Also in the educational context, the use of co-design refers to the process through which teachers, researchers, and developers, as well as students, collaborate to develop specific roles when it comes to creating teaching innovations (Roschelle et al., 2006). Its process can generate outcomes in the form of several prototypes. These outputs are also closely evaluated to ensure the effectiveness of such innovation, and to be able to enhance the positive outcomes of teacher instruction (Roschelle et al., 2006). Because the designed products are developed based on the ideas of multiple partners, the co-design process, engaging varied team members, can help to empower the implementation of interventions in the classroom. This means that such

* Corresponding author.
E-mail address: watcharasak.s@chula.ac.th (W. Sudla).
† Co-first authors.
E-mail address: chayut.p@chula.ac.th (C. Piromsombat).

https://doi.org/10.34044/j.kjss.2023.44.2.22
2452–3151/© 2023 Kasetsart University.

innovations are a close fit for the end-user, particularly teachers who are part of the design teams and play a critical role in implementing the proposed intervention for use in the classroom (Gravemeijer & van Eerde, 2009; Kwon et al., 2014; Lee, 2008; Matuk et al., 2016).

Co-design is a potential process that shifts teachers to collaborate with stakeholders and there are various empirical studies relating to the use of co-design to create interventions in many other countries, but there appears to be a lack of evidence that assembles or summarizes the advantages of co-design. Moreover, there is a limited understanding of what factors contribute to effective co-design. These viewpoints prompted the need for some investigations that can clearly summarize the knowledge of co-design. However, previous studies on co-design usually focused on student outcomes because the interventions were primarily aimed at improving student learning. To extend the understanding regarding the effectiveness of co-designed interventions on student outcomes, this study focused on the state of co-design in experimental research because these types of research focus on the manipulation of interventions to change the state of specific study settings and can measure the change in outcomes (Campbell & Stanley, 1963).

Based on the aforementioned situations, we developed the first research question, “How does co-design affect student outcomes?” To answer this, we aimed to describe the effect of interventions created through a co-design process on student outcomes. A meta-analysis was performed (Whittemore & Knafl, 2005) to bridge the gaps in our knowledge, which can provide answers to questions by analyzing the effect sizes of primary experimental and quasi-experimental studies. Since there was little evidence that described which factors influence the effectiveness of co-design, our second research question was, “How do moderating variables, defined as the study’s characteristics, influence the impact of co-design on student outcomes?” The second question may assist in gaining an understanding that researchers and practitioners can utilize when using co-design and addressing the question of how factors contribute to effective co-design. This question also led us to investigate possible moderators or potential study characteristics (i.e. type of intervention, education level, the explicitness of the co-design process, type of outcome measures, type of outcome variables, type of control group) on the effect of co-design on student outcomes.

Co-design is extensively implemented in empirical research, but the majority of them are non-experimental studies. Therefore, the number of robust experimental studies on co-design in education is still limited. Previous studies also revealed a low frequency of mutually-outcome variables between pieces of research. Therefore, this study considered the effect of co-design in teaching and learning on non-specific outcomes that cover the broad aspects of students’ cognitive, affective, and psychomotor domains, and all included studies had to use the common theme of co-design.

The principal purposes of this study were: (1) to explore the effect sizes of the learning activities used in co-design intervention on student outcomes; and (2) to analyze what factors influence such effects of co-design on student outcomes. A study that focuses on summarizing these issues could be helpful for educational research and practices in Thailand. Also, when it comes to using the notion of co-design in educational research with regard to further opportunities or specific aims, the current study’s findings may provide knowledge and methodological information and have useful results for stakeholders.

Methodology

Search Strategy

Seven academic databases in the form of Scopus, ERIC, Wiley, ScienceDirect, Web of Knowledge (ISI), and JSTOR were used to conduct the search. A manual search using search engines such as Google Scholar and a specific journal website devoted to co-design (i.e., CoDesign) were performed. A combination of search terms in Boolean search strings for co-design were identified by using PICO model (Davies, 2011). Participant (P) refers to teachers, instructors, and lecturers at various educational levels, as well as pre-service teachers. Intervention (I) refers to learning activities that apply or use the co-design process as the core approach for designing the intervention. Comparison (C) relates to those participants who were not in receipt of the co-design intervention or were not part of the experimental group. Instead, they received other forms of non-major intervention, such as competitive treatments, did not receive any interventions, or were taught using traditional methods. Outcomes (O) were identified as student outcomes only, because many researchers did not measure teacher outcomes. Even though teachers were identified as co-designers, this study only specified the scope of the outcome variables relating to students, including learning outcomes, competencies, performance, behavior, and the effectiveness of students. Previous studies, furthermore, used different terms for co-design, because the related words used had common roots in the form of participatory design, collaborative design, co-operative design, and co-creation (Zamenopoulos & Alexiou, 2018).

Inclusion and Exclusion Criteria

To conduct a meta-analysis in this study, we defined co-design as the process by which teachers collaborate with researchers, designers, and developers to create an educational innovation and implement these co-designed interventions to enhance student learning in classrooms (Roschelle et al., 2006). Regardless of how many phases of co-design the researchers used, the procedure included in this study included broad steps of understanding, planning, and implementing interventions in the classroom. Although some scholars and practitioners have introduced co-design under different terminologies (e.g., co-creation, collaborative design, and co-production) (De Koning et al., 2016; Koskela-Huotari et al., 2013), when conducting a review of the pieces of research, we used these terms interchangeably and with the same meaning.

To be included in the present study, the primary studies had to: (1) be empirical research peer-reviewed publications; (2) use an experimental or quasi-experimental research design with experimental and/or control groups–the experimental group was the one which received or achieved the learning activities that were created by using the co-design process; (3) ensure that the co-design process used in the primary research met the characteristics of co-design whether or not the sources used co-design or different terms (e.g. collaborative design, co-creation, partnership, design team). The included studies needed to demonstrate the process of co-design, in which a teacher was one of the co-designers and students were participants in the designed-treatment implementation. As well, the number of co-design steps used in the primary research were not constrained as to whether they used four steps of co-design; (4) be published in the English language because there is no empirical research in the Thailand context; (5) be research in education; and (6) be published between January 1, 2012, and January 20, 2022 (in-press included), over 10 years of publications. In addition, the term-used of variables of student outcomes of publication were not restricted. To be excluded from the present study, studies had to: (1) have been published as proceeding papers in an academic conference or as theses ("grey" literature was not included); (2) not provide adequate information with regard to conducting the analyses; and (3) have an unclear co-design or collaborative design process used in the paper, and its process not meeting the characteristic of co-design.

After the initial search, 259 records were obtained from the above-mentioned databases and other methods. By applying the inclusion and exclusion criteria, the recorded publications were filtered by using the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach (Page et al., 2021). Duplicate entries were removed, resulting in a total of 190 records. These records were filtered by considering the research abstract and title. A number of inappropriate records which did not meet our specifications were found; for instance, there were general academic papers which had inconsistent co-design content in terms of our specifications. We excluded unsatisfactory records, which left 101 articles, of which eight were found to be eligible. Figure 1 shows the PRISMA flow diagram for our search, identification, screening, and inclusion procedure.

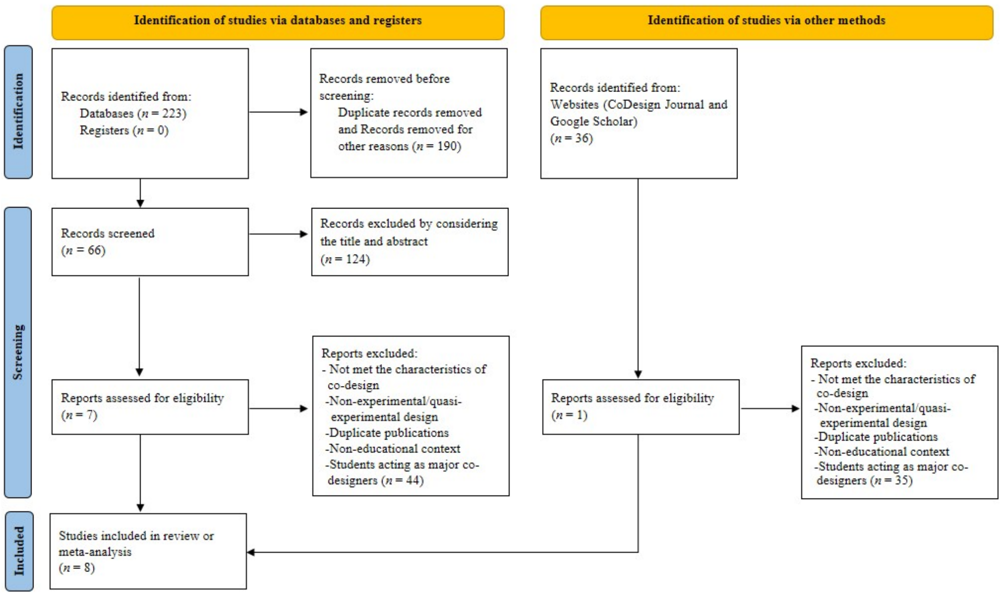

Figure 1 PRISMA flow diagram of study procedure

Coding Procedure

Our study identified the outline and coded the information for the eligible studies into four sections: (1) a description of the study (i.e. authors, study sample, etc.); (2) a description of the intervention (e.g. name of intervention, characteristics of the co-design process); (3) a description of the study design (i.e. type of outcome measures, education level of students, characteristics of the control group, etc.); and (4) the information required to calculate the effect size (i.e. sample means, standard deviations, etc.). The moderating variables in this study were not pre-specified, but their characteristics were coded in order to investigate them further in a subsequent process. This procedure was introduced to guide the researchers through the coding process. Two researchers were coders for the primary studies. The first coder coded the information from the studies, and the second coder checked and discussed this preliminary coding to resolve whether or not the coding process was acceptable.

Risk of Bias Assessment

Two researchers independently assessed the quality of the included studies by using Cochrane's risk of bias in non-randomized studies of interventions (ROBINS-I) (Sterne et al., 2019). The ROBINS-I tool was used to evaluate every aspect of the experimental studies in the form of pre-experiment, experiment, and post-experiment. The study quality and risk of bias were explored using seven dimensions: (1) bias due to confounding; (2) bias in terms of participant selection; (3) bias in the classification of interventions; (4) bias due to deviations from intended interventions; (5) bias due to missing data; (6) bias in outcome measurement; and (7) bias in the selection of the reported result. Each study under consideration was classified into four levels in the form of low risk of bias, moderate risk of bias, serious risk of bias, critical risk of bias, or, in the event of unclear information for judging purposes, it was classified as no information. The risk of bias assessment was guided by the Cochrane handbook (Sterne et al., 2019). We then used the "robvis" package in R version 4.0.5 to present or visualize the risk of bias (McGuinness & Higgins, 2020).

Publication Bias Assessment

Even though the number of primary studies in the current analysis was rather modest, the collected effect size was deemed sufficient. Two methodologies were used to assess publication bias. The first technique was the funnel plot, which is a scatter plot that shows the relationship between effect size and its standard error. Assuming that the distribution of effect size is symmetrical and appears consistently around the 95 percent confidence intervals (CIs), this result implies that there is no publication bias in the included studies (Sterne et al., 2011). Egger's regression test was the second technique for analyzing publication bias. This method is a statistical approach for detecting funnel plot asymmetry. If the test result is statistically significant, this indicates publication bias (Ferguson & Brannick, 2012; Van Aert et al., 2019). The publication bias assessment was carried out, and the assessment results were visualized by using the "funnel" and the "regtest" functions of the "metafor" package in R (Viechtbauer, 2010).

Data Analysis

The study characteristics were separated into two categories: context variables and methodological variables (i.e. the characteristics of participants, education level of students, the characteristics of outcome measures, research design, etc.). They were used to explain effect sizes of the co-design on student outcomes. Descriptive statistics (i.e., mean, standard deviation, and study sample size) were used to compute the corrected standardized mean difference (SMD or Hedges' g), which was suitably adjusted for meta-analysis. To investigate the dispersion of the effect size, the "metafor" package in R (Viechtbauer, 2010) was used to generate a forest plot.

A test for heterogeneity of effect size among the studies was investigated using Cochran's Q-test and I^2. Heterogeneity indicates the variation of effect size across a body of studies. The Q-test was defined as the weight squared deviations of the summary effect size (Del Re, 2015). If the Q-test outcome satisfies the degree of heterogeneity, it illustrates the inconsistency in study outcomes between studies (Del Re, 2015). I^2 is a statistical index or percentage. It provides information about the degree of variation that explains the extent of heterogeneity. An I^2 of 25 percent, 50 percent, 75 percent describes low, medium, and large degrees of heterogeneity respectively (Del Re, 2015). This study used a random effects pooling method because the different methods and sample characteristics across studies were likely to introduce deviation among the true effect sizes (Del Re, 2015). Subgroup and moderator analysis were performed to examine the inconsistency of effect size between study contexts that determines the afore-mentioned moderating variables. We used the "metafor" package in R version 4.0.5 to investigate heterogeneity, and to conduct subgroup analysis using the random effect method (Viechtbauer, 2010).

Results

Our findings were divided into two sections. First, we reported the study characteristics, testing for heterogeneity of effect size, risk of bias assessment, and testing for publication bias, which addressed the first research question. To address the second research question, we reported the results of subgroup analysis and potential moderator analysis.

RQ1: How Does Co-design Affect Student Outcomes?

1.1 Description of study characteristics and test for heterogeneity of effect size

Eight of 259 empirical studies met the inclusion criteria and were then assembled in a meta-analysis. The included studies were published from 2015 to the first quarter of 2022. Most of the participants were university students, the number of which ranged from 30 to 253 students. The main stakeholders in the co-design process consisted of both teachers and students. The teachers were co-designers, whereas the students were the experimental units and acted as the units of analysis for measuring the change in outcome. By computing the Hedges' *g* or corrected SMD, 17 effect sizes from the included studies were identified. In terms of eligible studies, four of the studies produced co-design products in the form of artifacts which were tangible innovations (i.e. educational tools, equipment, devices) for teaching and learning. Another four of the studies delivered co-design products in the form of instructional methods or teaching approaches for use by teachers or university instructors.

The included studies showed several methods for measuring student outcomes. Four of the studies measured knowledge (i.e. achievement, concepts, understanding, etc.), whereas five of the studies measured student skills (i.e. writing performance, reasoning skills, etc.). Knowledge and skills tended to be measured through the use of tests, but some studies were measured with the help of questionnaires. Applications of the co-design process in the research process was considered by the clarity of the concepts presented in the original articles. If the co-design process used in the included studies was clearly defined, or the process met with the core concept of co-design, those studies were categorized as an explicit group. Based on our specification, six of the studies were explicit co-designs, while two of the studies were classified in the implicit group; this was because the studies identified or mentioned the co-design concept, but the process was still ambiguous. One of the studies used an active control group whereas seven of the studies used an inactive control group which covered no treatment and competitive treatments. Table 1 shows a summary of the characteristic of the eligible studies.

This synthesis led to the identification of 17 effect sizes from eight primary studies. Effect sizes were estimated by Hedges' *g* (SMD), and ranged from 0.41 to 2.08. The heterogeneity test showed that the effect size of the primary studies was heterogeneous ($Q(16) = 54.67$, $p < .001$). Consistently, there was a moderate to high degree of heterogeneity among the studies ($I^2 = 67.92\%$). The random-effects (REML) model was first used and yielded a pooled effect size; however, due to the high heterogeneous effect size across the included studies, it could not be used to assess the overall effect size. Initially, the result revealed that there is a high effect of co-design on overall student outcomes, which cover knowledge and skills ($k = 17$, Hedges' $g = 0.96$, 95% CI [0.75, 1.17], $Z = 9.06$, $p < .001$) (Cohen, 1988). Figure 2 shows the forest plot which describes the estimated effect size distribution.

Table 1 Characteristics of primary studies included in the analysis

Author (Year)	*n*	Education level of student	*k*	Study Characteristics				
				Type of co-designed product	Domain of outcome variable	Type of outcome measure	The explicitness of co-design step	Type of control group
Cviko et al. (2015)	105	Elementary	1	Artifact	Skills	Questionnaire	Explicit	Inactive
Shein & Tsai (2015)	69	Secondary	3	Instruction	Skills	Test & Questionnaire	Explicit	Inactive
Kyza & Nicolaodou (2017)	68	Secondary	2	Artifact	Knowledge, Skills	Test	Explicit	Inactive
Wang (2017)	30	Secondary	3	Artifact	Knowledge	Test	Explicit	Inactive
Lui et al. (2020)	234	Higher	3	Instruction	Knowledge	Test	Implicit	Inactive
Sanina et al. (2020)	253	Higher	3	Artifact	Skills	Test	Explicit	Active
Saka et al. (2021)	135	Secondary	1	Instruction	Knowledge	Test	Implicit	Inactive
Fu et al. (2022)	100	Higher	1	Instruction	Skills	Test	Explicit	Inactive

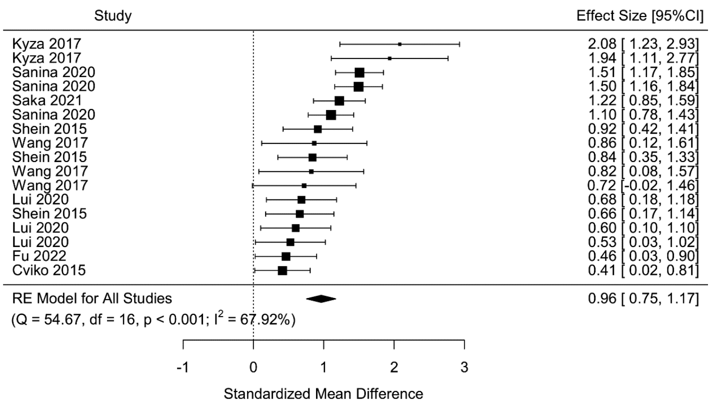

Figure 2 The dispersion of estimated effect sizes, confidence intervals, and overall effect size of co-designs on student outcomes

1.2 Risk of bias assessment

The risk of bias assessment revealed that most of the primary studies had rarely information about deviation from the intended interventions and reporting the missing data. Additionally, the assessment results highlighted that most of the primary studies had an overall moderate to serious risk of bias due to the selection of the participants, because they have rarely provided reasons for random assignment. Most of the studies — 70 percent to 75 percent — were evaluated as having a low risk of bias in measurement of outcomes because they rarely reported crucial information regarding the validity evidence of measurements. As well as bias due to the selection of the reported results, some studies still have not reported important statistics. Some studies had a low to moderate risk of bias due to confounding variables because there was seldom information about controlling the extraneous factors. Figure 3 visualizes the assessment of risk of bias by the researchers.

1.3 Publication bias assessment

The result indicated a symmetrical distribution around the effect size with a few outliers (see Figure 4). This suggests a less likely existence of publication bias. The Egger regression test of asymmetry of the funnel plot was used to objectively investigate publication bias. The findings indicated that there was no confirmation of publication bias ($t(15) = -0.42$, $p = .677$).

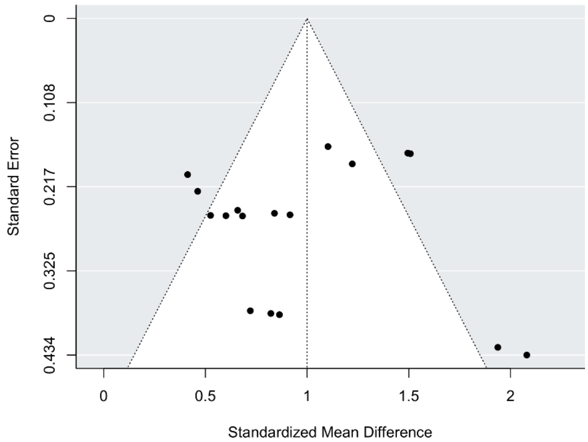

Figure 4 Relationship of estimated effect size and its standard error

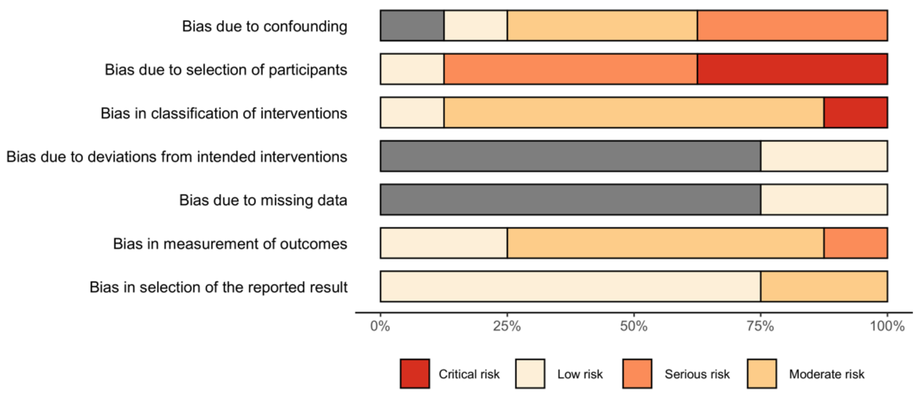

Figure 3 Summary of risk of bias assessment using Cochran's ROBINS-I tools

RQ2: How Do Moderating Variables Influence the Impact of Co-Design on Student outcomes?

Subgroup and moderator analyses

Due to the heterogeneity test, the results showed heterogeneous effect sizes of co-designed interventions on student outcomes. We then used the random effect model for exploring the role of the moderating variables to explain the difference between those effect sizes. Subgroup analysis was conducted to identify whether or not additional moderators could account for the variance observed in the samples. Based on the RE Model, the co-design product is an essential factor which indicates a significant difference between studies because the effect size between two types of co-designed products was statistically significant ($Q(2) = 98.24$, $p < .001$) and had very high level of heterogeneity across effect sizes ($Q(15) = 41.99$, $p < .001$, $I^2 = 62.81\%$). Therefore, we used the co-designed product as the major moderator for subgroup analysis. The co-design products in this study were identified as belonging to the "Instructional methods group" or "Instruction group" and to the "Artifacts group". The Instruction group takes the form of products emerging from the co-design process such as instructional methods, principles, and teacher manuals for managing instruction to students, whereas the Artifacts group takes the form of products such as tangle interventions in the form of media, tools, and equipment which the teacher used in their teaching practices. There were four studies classified to the Artifact group, whereas later four studies were classified to the Instruction group. The results are shown as a forest plot in Figure 5.

With regard to each subgroup, the findings highlighted that there was still heterogeneity in terms of the effect sizes in some subgroups. The heterogeneity test showed that the effect size of the Instruction group was homogeneous (Q (7) = 9.83, p = .199) and there was low heterogeneity among the studies ($I^2 = 33.86\%$). Due to the heterogenous effect size in the instruction group, the result revealed that there is a moderate effect of co-design on overall student outcomes (k = 8, Hedges' g = 0.76, 95%CI [0.55, 0.96], Z = 7.31, p < .001) (Cohen, 1988). On the contrary, the heterogeneity test pointed out that the effect size of the Artifact group was heterogeneous ($Q(8) = 32.16$, $p < .001$), and there was high heterogeneity among the studies ($I^2 = 76.51\%$). This evidence indicates that there was still a degree of variation among the primary studies in the Artifacts group. That is, the effect size of the co-design products in the form of artifacts on student outcomes could be explained by other moderators. Prior to analyzing moderator analysis, however, the overall RE model revealed that the Artifact group had a high effect size (k = 9, Hedges' g = 1.19, 95%CI [0.84, 1.54], Z = 6.67, p < .001) (Cohen, 1988).

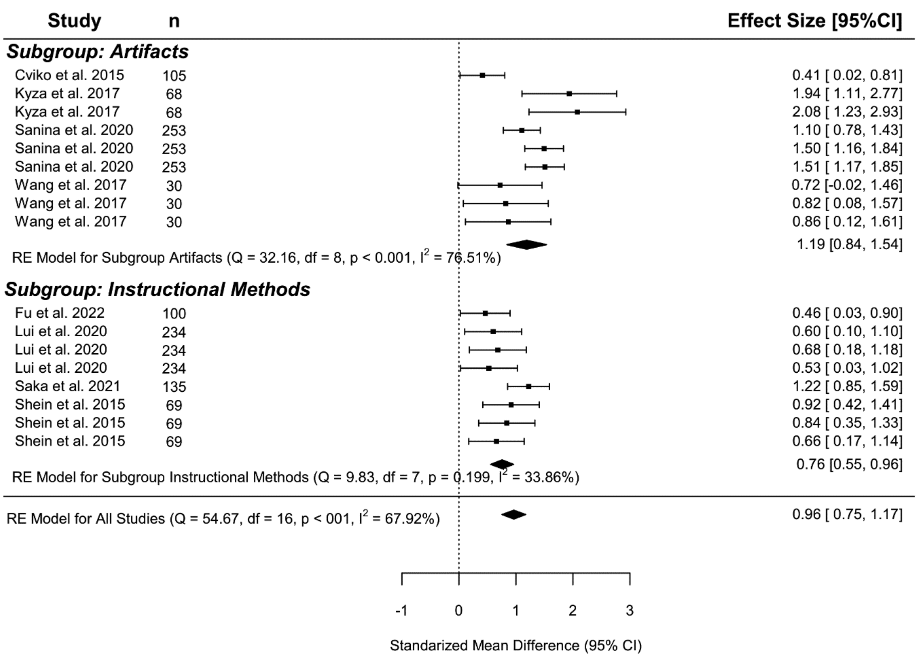

Figure 5 Subgroup analysis results in terms of estimated effect sizes, confidence intervals, and the overall effect size of co-design on student outcomes

Table 2 shows results of the moderator analysis. We investigated the difference of effect size in the Artifact group. Five moderators covering education level of students, the explicitness of the co-design process used in the study, type of outcome measures, the domains of outcome variables, and the type of control group, were carried out only for the Artifact subgroups. The moderator analysis demonstrates that the education level significantly moderated the effects of the co-designed intervention on student outcomes of the Artifact group ($Q(3) = 73.84$, $p < .001$). In elementary, primary, and secondary school, the co-design products in the form of artifacts were more likely suited to the students in secondary and university compared with elementary and primary students. Owing to the limited number of studies in the Artifact group moderator analyses could not be undertaken for the Artifact group in terms of the explicitness of co-design principles and type of outcome measures.

In addition, the results revealed that the domains of the outcome variables significantly moderated the effects of the co-designed intervention on student outcomes in the Artifact group ($Q(2) = 39.69$, $p < .001$). That is, students who took a measurement in the skills had more effective outcomes than the knowledge (concepts, understanding, etc.). Finally, in terms of the types of the control group as a moderator, the researchers who assigned students into the active control group had more effective outcomes than did those in the inactive group ($Q(2) = 40.12$, $p < .001$).

Our findings revealed that co-design has a heterogeneous effect on student outcomes and differs in various conditions of experiments, especially co-designed products. Also, there are potential moderators to explain variation of these effects. These will be discussed in the following section.

Discussion

Overall, a meta-analysis of eight eligible studies with 17 effect sizes indicated that the effect sizes differed in terms of characteristics, along with the research design of the primary studies. According to this explorative study, the characteristics that moderated those effects consisted of six variables (i.e. type of co-designed products, education level of students, the explicitness of the co-design process, type of outcome measures, nature of outcome variables, and type of control group). Based on the findings, some additional ideas that could be discussed are as follows:

The results demonstrated a heterogeneous effect size among the included studies. A meta-analysis researcher should therefore carefully combine the effect size with a random-effects model (Del Re, 2015). The effect sizes perhaps deviated when the learning activities were conducted in several different contexts. In this study, the learning activities that used co-designed products as artifacts likely had a higher impact on student outcomes than instructional methods. The explanation for this finding could be that the co-designed products, in this case, were

Table 2 Results of the moderator analyses for co-design on student outcomes in the artifact subgroup

Moderators	*k*	Hedges' *g*	*SE*	*Z*	95%CI	Q	*df*
Education level of student							
Elementary & Primary school	1	0.41	0.38	1.10	[-0.33, 1.15]	73.84***	3
Secondary school	5	1.24	0.23	5.44***	[0.79, 1.69]		
Higher education	3	1.37	0.21	6.56***	[0.96, 1.76]		
The explicitness of co-design step							
Explicit	9	1.19	0.18	6.67***	[0.84, 1.54]	32.16***	8
Unclear							
Type of outcome measures							
Test	9	1.19	0.18	6.67***	[0.84, 1.54]	32.16***	8
Questionnaire							
Domain of outcomes variables							
Knowledge	4	1.10	0.31	3.52***	[0.49, 1.71]	39.69***	2
Skills	5	1.25	0.24	5.22***	[0.78, 1.72]		
Type of control group							
Active	3	1.37	0.27	5.01***	[0.83, 1.90]	45.72***	2
Inactive	6	1.06	0.23	4.54***	[0.60, 1.51]		

Note. Each moderator was separately tested. However, the overall mixed-effects model test for residual heterogeneity based on all moderators was insignificant at level of .05 $Q(5) = 10.75$, $p = .057$.
*** $p < .001$.

tangible innovations. The concrete interventions presumably support the learning of students in the form of authentic participation. It is possible that students can contact and interact with the real equipment, which could further encourage the learning and improve the affective behavior of students. Moreover, interaction with tangible artifacts potentially promotes higher outcomes for students (Erstad, 2002; Rountree et al., 2002).

In terms of the moderator analysis results by subgroup, we discovered that there is little variation across effect sizes in the instructional group, but high variation in the artifact group. In particular, the findings indicate that the higher education students who were involved in the artifact group seemed to have higher outcomes than the other groups. We could clarify this finding by suggesting that in the case of the inexperienced students, a novice group, when they interact with a complex innovation, the outcomes are less favorable. This indicates that the complex innovations were probably more suitable for students who had more learning experience (Rountree et al., 2002). The research findings show that the students who participated in the learning activities involving artifacts had better outcomes than the other groups. This is because the nature of the co-design process emphasizes the product in the case of tangible innovations, in that the process begins with brainstorming the idea and transferring it to the various prototypes (Burkett, 2019; Easterday et al., 2018). Tangible co-design products may yield more effective outcomes than intangible innovations.

The domains of outcome variables showed a difference in the effects on student outcomes in the artifact group. Thus, researchers should pay particular attention to selecting an appropriate domain of outcome variables linked to the nature of the designed intervention. Within the artifact group, the research findings illustrate that the student outcomes in terms of skills would be higher than those of knowledge. This is because students who interact with tangible innovations likely participate in physical practices. Learning that involves such practices may increase student skills in conjunction with learning content (Escobedo et al., 2012).

Finally, with regard to the type of control group, the active control group had better outcomes in the artifact group. This is because active control groups are frequently used in experimental design in educational research, given that such an approach allows careful control of extraneous variables, whereas the use of inactive control groups as found generally in a quasi-experimental design, is more susceptible to control of confounding variables (Campbell & Stanley, 1963; Street, 1995).

The present study focused on a synthesis of experimental research that has been applied to the co-design principle, even if the number of such studies is a limited number of publications in educational research. Mostly, the consideration of co-design on the part of educational researchers is frequently related to non-quantitative data, and rarely uses a stringent experimental design (i.e. design-based research, design research, etc.). Researchers commonly collect developmental and qualitative data in order to explain the outcomes from the point of view of the participants. More research will be necessary to develop a research synthesis that is not based on experimental research and on effect sizes. Hopefully, further research syntheses might support the strength of the co-design methodology in education.

Conclusion and Recommendation

Even though several scholars have suggested that the co-design process can promote effectiveness in education, the intervention which constructs based on co-design has different effects on the target outcomes in manifold contexts. Thus, to apply the idea of co-design, practitioners should concern themselves with some of the confounding variables that would interfere with the desired outcomes. Although the findings of this study provide knowledge of co-design on student outcomes, there has been hardly any co-design research in education in Thailand. Education sectors should cooperate with relevant organizations to support the use of the co-design methodology in education. This furtherance could possibly increase the importance of co-design in educational research in Thailand and strengthen the wide application of co-design methodology.

Based on the present study, we did understand that there was a very small number of studies included for this meta-analysis, but the information synthesized through the primary research might provide a lesson learned with regard to co-design in education. Even though there were some limitations about drawing a conclusion regarding the effectiveness of co-design, we strictly considered and followed the standards of a robust protocol for conducting a meta-analysis that could provide systematic evidence of co-design knowledge.

Conflict of Interest

The authors declare that there is no conflict of interest.

References

Burkett, I. (2019). *An introduction to co-design.* https://www.yacwa.org.au/wp-content/uploads/2016/09/An-Introduction-to-Co-Design-by-Ingrid-Burkett.pdf

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Houghton Mifflin.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Psychology Press.

Cviko, A., McKenney, S., & Voogt, J. (2014). Teacher roles in designing technology-rich learning activities for early literacy: A cross-case analysis. *Computers & Education*, *72*, 68–79. https://doi.org/10.1016/j.compedu.2013.10.014

Davies, K. S. (2011). Formulating the evidence based practice question: A review of the frameworks. *Evidence Based Library and Information Practice*, *6*(2), 75–80. https://doi.org/10.18438/B8WS5N

De Koning, J. I., Crul, M. R., & Wever, R. (2016, May 24–26). *Models of co-creation* [Paper presentation]. The ServDes. 2016 Conference, Copenhagen, Denmark. http://www.academia.edu/download/55012757/ecp16125022_small.pdf

Del Re, A. C. (2015). A practical tutorial on conducting meta-analysis in R. *The Quantitative Methods for Psychology*, *11*(1), 37–50. https://doi.org/10.20982/tqmp.11.1.p037

Easterday, M. W., Rees Lewis, D. G., & Gerber, E. M. (2018). The logic of design research. *Learning: Research and Practice*, *4*(2), 131–160. https://doi.org/10.1080/23735082.2017.1286367

Erstad, O. (2002). Norwegian students using digital artifacts in project-based learning. *Journal of Computer Assisted Learning*, *18*(4), 427–437. https://doi.org/10.1046/j.0266-4909.2002.00254.x

Escobedo, L., Nguyen D. H., Boyd, L., Hirano, S., Rangel, A., Garcia-Rosas, D., … Hayes, G. (2012, May 5–10). *MOSOCO: A mobile assistive tool to support children with autism practicing social skills in real-life situations* [Paper presentation]. The SIGCHI Conference on Human Factors in Computing Systems, Austin, TX. https://doi.org/10.1145/2207676.2208649

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods*, *17*(1), 120–128. https://doi.org/10.1037/a0024445.

Fu, Q. K., Zou, D., Xie, H., Cheng, G., & Hwang, G. J. (2022). Effects of a collaborative design approach on pre-service teachers' ability of designing for learning with a digital game. *Education and Information Technologies*, *27*, 1–24. https://doi.org/10.1007/s10639-021-10818-3

Gravemeijer, K. & van Eerde, D. (2009). Design research as a means for building a knowledge base for teachers and teaching in mathematics education. *The Elementary School Journal*, *109*(5), 510–524. https://doi.org/10.1086/596999Koskela-Huotari, K., Friedrich, P., & Isomursu, M. (2013, June 18–21). *Jungle of "co"* [Paper presentation]. The Naples Forum on Service, Napoli, Italy. https://pdfs.semanticscholar.org/ec0f76d775bc4496683db0d.pdf

Kwon, S. M., Wardrip, P. S., & Gomez, L. M. (2014). Co-design of interdisciplinary projects as a mechanism for school capacity growth. *Improving Schools*, *17*(1), 54–71. https://doi.org/10.1177/1365480213519517

Kyza, E. A., & Nicolaidou, I. (2017). Co-designing reform-based online inquiry learning environments as a situated approach to teachers' professional development. *CoDesign*, *13*(4), 261–286. https://doi.org/10.1080/15710882.2016.1209528

Lee, Y. (2008). Design participation tactics: The challenges and new roles for designers in the co-design process. *Co-design*, *4*(1), 31–50. https://doi.org/10.1080/15710880701875613

Lui, M., McEwen, R., & Mullally, M. (2020). Immersive virtual reality for supporting complex scientific knowledge: Augmenting our understanding with physiological monitoring. *British Journal of Educational Technology*, *51*(6), 2181–2199. https://doi.org/10.1111/bjet.13022

Matuk, C., Gerard, L., Lim-Breitbart, J., & Linn, M. (2016). Gathering requirements for teacher tools: Strategies for empowering teachers through co-design. *Journal of Science Teacher Education*, *27*(1), 79–110. https://doi.org/10.1007/s10972-016-9459-2

McGuinness, L. A., & Higgins, J. P. T. (2020). Risk-of-bias visualization (robvis): An R package and shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods*, *12*(1), 55–61. https://doi.org/10.1002/jrsm.1411

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron I., Hoffmann T. C., Mulrow, C. D., … Moher, D. (2021) The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, *10*(1), 1–11. https://doi.org/10.1186/s13643-021-01626-4

Roschelle, J., Penuel, W., & Shechtman, N. (2006). Co-design of innovations with teachers: Definition and dynamics. *ICLS '96: Proceedings of the 1996 international conference on Learning sciences*, 606–612. https://dl.acm.org/doi/10.5555/1150034.1150122

Rountree, J., Wong, W., & Hannah, R. (2002). Learning to look: Real and virtual artifacts. *Journal of Educational Technology & Society*, *5*(1), 129–134. https://www.jstor.org/stable/pdf/jeductechsoci.5.1.129.pdf

Saka, O. A. (2021). Can teacher collaboration improve students' academic achievement in junior secondary mathematics? *Asian Journal of University Education*, *17*(1), 33–46. https://doi.org/10.24191/ajue.v17i1.8727

Sanders, E. B. N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *Co-design*, *4*(1), 5–18. https://doi.org/10.1080/15710880701875068

Sanina, A., Kutergina, E., & Balashov, A. (2020). The co-creative approach to digital simulation games in social science education. *Computers & education*, *149*, 103813. https://doi.org/10.1016/j.compedu.2020.103813

Shein, P. P., & Tsai, C. Y. (2015). Impact of a scientist–teacher collaborative model on students, teachers, and scientists. *International Journal of Science Education*, *37*(13), 2147–2169. https://doi.org/10.1080/09500693.2015.1068465

Sterne, J. A., Hernán, M. A., McAleenan, A., Reeves, B. C., & Higgins, J. P. (2019). Assessing risk of bias in a non-randomized study. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (pp. 621–641). https://doi.org/10.1002/9781119536604.ch25

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., ... Tetzlaff, J. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Bmj*, *343* https://doi.org/10.1136/bmj.d4002

Street, D. L. (1995). Controlling extraneous variables in experimental research: A research note. *Accounting Education*, *4*(2), 169–188. https://doi.org/10.1080/09639289500000020

Van Aert, R. C., Wicherts, J. M., & Van Assen, M. A. (2019). Publication bias examined in meta-analyses from psychology and medicine: A meta-meta-analysis. *PloS One*, *14*(4). https://doi.org/10.1371/journal.pone.0215052

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wang, Y. H. (2017). Exploring the effectiveness of integrating augmented reality-based materials to support writing activities. *Computers & Education*, *113*, 162–176. https://doi.org/10.1016/j.compedu.2017.04.013

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of advanced nursing*, *52*(5), 546–553. https://doi.org/10.1111/j.1365-2648.2005.03621.x

Zamenopoulos, T., & Alexiou, K. (2018). *Co-design as collaborative research. Bristol University/AHRC Connected Communities Programme.* https://connected-communities.org/wp-content/uploads/2018/07/Co-Design_SP.pdf