

Estimating Effect Sizes, Heterogeneity Parameters and Weighted Standard Deviation (WSD) of Postgraduate Theses using Meta-Analytic and Systematic Review Methods

Simon Ntumi*

Department of Educational Foundations, University of Education, Winneba (UEW), West Africa, Ghana

*Corresponding author: sntumi@uew.edu.gh

Received June 07, 2024; Revised July 08, 2024; Accepted July 15, 2024

Abstract The purpose of the study was to conduct a meta-analytic from an institutional repository by exploring the statistical analysis and reporting practices of postgraduate students' theses in the University of Cape Coast. To achieve this, the study was nested into the quantitative approach where archival data were retrieved from University of Cape Coast institutional repository (UCCIR). The 2020 PRISMA chart flow was used to extract 778 researched studies from the UCCIR. The study found overall medium resultant significant effect size indicating the extracted studies may have some limited practical applications (Min. $ES=.378$; $p=.000$; $Hg=1.72$, $z=12.20$; $Std\ Err=.238$; $n=752$; Max. $ES=.430$; $p=.012$; $Hg=.812$, $z=14.12$, $Std\ Err=.623$; Overall $=.591$, $p=.000^{**}$). Again, the results from the heterogeneity analysis also showed that there was some degree of probability sampling errors in the extracted studies. Also, the study found that most of the extracted studies were likely to produce similar results and conclusions as a result of similar statistical methods employed by the students ($T^2=.627$, $p=.023^{**}$, $df=6$, $Z=16.12$; $CI=95\%$, $n=752$; $Pq=.934$, $p=.002^{**}$, $df=6$, $Z=12.01$; $CI=95\%$; $Tau=.723$, $p=.004^{**}$, $df=6$, $Z=30.01$; $CI=95\%$; $n=752$; $WSD=.075$, $p=.013^{**}$, $df=6$, $Z=17.23$; $CI=95\%$; $n=752$). From the ensued findings, it was concluded that there were some statistical misapplications and misinterpretations of studies results leading to different conclusions and recommendations from the studies. The study recommended that to improve upon the robustness and rigorousness of postgraduate studies in Ghana, research mechanisms such as series and timely training workshops should be put in place by the university to re-orient and expose postgraduate students to modern statistical methods.

Keywords: statistical analysis, postgraduate thesis, meta-analysis, reporting practices, institutional repository

Cite This Article: Simon Ntumi, "Estimating Effect Sizes, Heterogeneity Parameters and Weighted Standard Deviation (WSD) of Postgraduate Theses Using Meta-Analytic and Systematic Review Methods." *American Journal of Applied Mathematics and Statistics*, vol. 12, no. 3 (2024): 41-54. doi: 10.12691/ajams-12-3-2.

1. Introduction

In the global statistical literature, the manual application of statistical software was very part of data analysis [1]. During these modern times of analysis, researchers struggled in giving meaningful interpretation to their data [2,3]. In essence, researchers are experiencing a transition from manual analysis with paper to more efficient digital/electronic analysis with statistical software [2]. Clearly, from Matthew and Sunday [2], it is posited that the emergence of modern statistical software has undoubtedly contributed enormously to the development in research studies in this 21st century.

In modern educational research, a variety of statistical methods and techniques are employed across different academic disciplines, based on the specific experimental data and analysis requirements of the researcher [4,5]. In educational research, there are some areas of statistical

methods that are applicable. These include; descriptive statistics, which deals with the grouping, tabulation, graphical representation and a quantitative description of the data using measures of the central tendencies [4,5].

Invariably, scholars have admonished that if all article authors who are competent in statistical methods would not publish poor statistical results and report those to the funding agencies or journal editors involved, then the recipients would become more motivated to enforce sound statistical practice to inform policies and decisions [6,7,8,9]. In clear terms, the researcher shares in the assertions that the research community should not let ethics of research and statistics sit unused on shelves or in unvisited cyberspace where researchers could publish any results to society [6,10,11,12]. Undoubtedly, if statistical assumptions are not appropriately considered during the selection of statistical tests, statistical misapplications and misinterpretation of results are likely to occur [1,11,13]. However, this comes with a great deal of work that needs to be mitigated and suppressed within the space of

educational literature. Therefore, in line with the above empirical and statistical assumptions, this study is therefore conceptualised and occasioned on these concerns and worrying phenomena in the statistical world of educational research.

Nesting the above propositions into a plethora of researched works on misuse and abuse of statistical analysis and reporting, one striking idea that is noted is that these growing concerns (statistical misapplication) are not only limited and found in local studies; however, it appears to be a global cancer that travels a long way to affect most decisions and conclusions on national and global issues [8,14]. More specifically, from many statistical works and reviews, it is evident that reporting and analysis of statistical assumptions or models are sometimes absent from many research articles and theses [8,14,15].

To empirically support and validate these observations and experiences, several related studies have provided empirical evidences. For example, in the study of Bland and Altman [16], it is reported that some published articles and theses evaluated lacked any discussion of the statistical assumptions and conditions they employed. Sequel to the above, a review from Ntumi and Twum Antwi-Agyakwa [15] suggest that some studies may have inconsistent analysis and as such, their conclusions and recommendations may be inconsistent and problematic due to lack of reporting psychometric properties (e.g. validity and reliability).

In their investigation, Ntumi and Twum Antwi-Agyakwa [15] analysed 763 articles published between 2010 and 2020, sourced from 15 education-related journals published by five different publishing houses. The study uncovered that over half of the articles reviewed failed to include statistics pertaining to either validity ($n=456$ out of 763, 59.8%) or reliability ($n=400$ out of 763, 52.4%). Among those that did report validity or reliability statistics, the Alpha Coefficient was the most frequently used measure to establish reliability ($n=185$, 50.9%), while face validity was commonly cited ($n=219$, 71.3%) to establish validity. This result suggested that some studies may not have given much attention to psychometric properties such as validity and reliability which have some resultant consequences on findings and recommendations.

Similarly, Ghasemi and Zahediasl [17] reported that statistical misapplications are common in scientific literature and about 50% of the published articles have at least one error and this could affect the conclusions of the data. They emphasised and projected that the assumption of normality and other related assumption need to be checked for many statistical procedures, namely parametric tests, because their validity depends on it. All these empirical propositions founded and grounded the exegesis of this study. Akin to the above, Sabharwal [11] also reported that selecting the inappropriate statistical analysis and sample size is a very common problem for postgraduate students. The more worrying phenomenon is that many articles and theses fail to report which statistical tests were utilised during data analysis and the underlying assumptions for these statistical tools [11,18]. Only stating that tests were used “where appropriate” is also grossly inadequate, yet commonly done [11,18].

These narratives suggest that if research students (including postgraduate students at the University of Cape Coast) are not much robust and rigorous enough in their use of statistical tools and models, it is highly possible that the decisions and inferences from their results could adversely harm and costs society [7,13]. Evidently, poor statistics in science may lead to insufficient scientific evidence and conclusions, likewise poor statistics lead to poor social sciences policies, decisions and programmes. For all of these consequences and structural effects, it is important that graduate students conducting research in different capacities pay careful attention to the quality of statistical analysis and reporting in their areas of jurisdiction and competence [9,11,13].

Narrowing the argument to more specific and practical terms, the researcher has witnessed several theses defence and academic conferences at different levels of the academic ladder. In some cases, at the defence and conference presentations, plethora evidences suggest that some of the students or presenters seem to lack some basic rudiments and assumptions for the choice of their statistical tools and the interpretation thereof. These assertions validate the conclusions of Ntumi and Twum Antwi-Agyakwa [15], who found that some educational studies abuse some rudiments and assumptions for the choice of statistical tools, and this eventually affects the conclusions and the interpretations of those studies.

In all these expounded issues on statistical tools, the impression created that, it is highly possible and evident that some statistical misapplications (analysis) and misinterpretations could be found among some social science research works (MPhil, published articles and PhD theses) conducted among graduate students in the University of Cape Coast. When these observations turn true, it will go to confirm the studies of Mishra et al. [13]; Matthew et al, [18] and Sabharwal [11] who akin asserted that graduate students research works contain some statistical misapplications and these errors affect their conclusions and recommendations thereby leading to severe consequences on policy and decision making.

The formation and the conceptualisation of the study is based on a number of identified research gaps. The research gaps leading to this study are presented in the following paragraphs. From the theoretical perspective, it appears that issues of statistical analysis and reporting using a meta-analysis and systematic review are completely missing in the local literature, especially within the space of measurement and evaluation [15,19]. Clearly, the application of statistical decision theory (SDT) in statistical analysis and reporting is very rare in both local and international literature.

Arguably, it looks like this theory (statistical decision theory) has not be fully conceptualised and framed in research studies especially in the Ghanaian literature [19,20]. This underutilisation may be due to an incomplete understanding of the conceptual underpinnings of statistical decision theory. To bridge this gap, this study is needful to review and expand statistical decision theory (SDT) in line with meta-analysis and systematic review and how statistical analysis could be reported to reduce or minimise statistical misapplications in research works to help make informed statistical and policy decisions.

In the international literature, it appears that meta-analytic and systematic reviews as research approaches have been well documented and represented [21]. Due to their relevance, these approaches have helped to track research approaches that could help produce refined and evidence-based findings. Tawfik et al. [21] highlighted that systematic reviews and meta-analysis (SR/MAs) are situated at the top of the evidence-based pyramid, indicating a high level of evidence. As such, a well-executed SR/MA is considered a viable approach for keeping clinicians abreast of current evidence-based research in various fields of study, including education.

For example, in the field of education, scholars such as Kelley et al. [22], Moreno-Peral et al. [23], Fry et al. [24], Sohn et al. [10], Garzón et al. [25] have used systematic review and meta-analysis (SR/MA) in exploring several issues including the selection of appropriate statistical methods. According to the research of Moreno-Peral et al. [23], it was discovered that psychological and/or educational interventions are efficacious in the prevention of anxiety, with a statistically significant but small overall effect size. The aforementioned conclusion was based on the analysis of 29 randomised controlled trials (36 comparisons) that encompassed a total of 10,430 patients across 11 different countries spanning four continents. Through sensitivity analyses and adjustment for publication bias, it was determined that the overall effect size was resilient. It is important to note that substantial heterogeneity was observed, which was accounted for by a meta-regression model. Clearly, most of these international studies have concluded and proven that systematic review and meta-analysis (SR/MAs) are very essential methods for synthesising quantitative results of different empirical studies.

Relatedly, most of authors have demonstrated that how meta-analysis are used in estimating their effect sizes (ESs) using Cohen's d and Hedge's g approach [24]. Others have used meta-analysis to estimate the heterogeneity parameters (Q -statistic, I^2 , T^2 , P_q , τ) and weighted standard deviation [10]. Using meta-analysis to estimate variance components (VCs) and standard errors (Std Errs) of students is evident in Moreno-Peral et al. [23]. Using meta-analysis to estimate random and fixed effect models is well documented in the work of Moreno-Peral et al. [23]. Estimating meta-regression coefficients of studies using the random effect model and method of moments estimation approach is also evident in the study of Rigabert et al. [26].

However, within the trail of Ghanaian literature, it seems that meta-analysis and systematic review as research approach is not really common [27]. These approaches have not been given the needed attention and exposure especially within the framework of Ghanaian literature. Searching through the research databases (the University of Cape Coast institutional repository and other tertiary repositories in Ghana), the researcher rarely came across any study that used meta-analysis and systematic review as a methodological approach. This study could serve as a base line to bridge the research approach gap and expand knowledge in meta-analysis and systematic review in the Ghanaian literature and beyond.

Coupled with the above identified gaps, the researcher's experiences at several research conferences, seminars,

proposals and theses defences have also exposed me to some structural and statistical misapplications and misinterpretations in some studies. Again, the researcher's interpersonal interactions with many postgraduate students in some public universities in Ghana may confirm the fact that some studies may ride on statistical misapplications and assumptions that may lead to wrong interpretations and invalid conclusions. All these observations may appear invalid until empirical evidence is established and documented. Unfortunately, however, searching through the Ghanaian literature (universities and colleges repository and other Ghanaian publication houses or hubs), the researcher could not fall on any empirical evidence on meta-analysis and systematic review to support these claims [15]. This has therefore created an empirical gap in the Ghanaian literature that needs attention and consideration. This study is therefore, formulated and conceptualised to provide localised empirical evidence on meta-analysis and a systematic review of how statistical analyses are reported and interpreted within the Ghanaian context.

Based on the empirical and theoretical evidences provided in the preceding paragraphs, hypothetically, wrong choice of statistical tool and their interpretations may not be farfetched from most Ghanaian institutional repositories especially at the University of Cape Coast. In response to these phenomena and the research gaps provided, a thorough and comprehensive meta-analysis and systematic review from the University of Cape Coast institutional repository (UCCIR) could be helpful. This may help track and establish empirical evidence and theoretical groundings to increase objectivity and establish more complete and accurate results of postgraduate studies.

Postgraduate theses play a critical role in advancing knowledge in various fields of study. However, the lack of consistency and rigour in statistical analysis and reporting practices employed in these theses can lead to inaccurate or incomplete interpretations of research findings. This problem may compromise the validity and generalizability of research, and limit the impact of research findings. This will help identify potential areas of improvement, and develop strategies to enhance the rigour and consistency of these practices. Hence, the need and the justification for this study.

2. Research Questions

1. What are the combined estimated effect sizes of postgraduate theses deposited at UCCIR using Cohen's d and Hedge's g approach?
2. What are the combined estimated heterogeneity parameters (Q -statistic, I^2 , T^2 , P_q , τ) and weighted standard deviation (WSD) of postgraduate theses deposited at the UCCIR?

3. Theoretical Review

Statistical Decision Theory (SDT)

To gain sound theoretical groundings and implications for a study, it is said that there is nothing more practical than a good theory or a model. To place the conceptualisation of the study in theoretical lens, the study

adopted the statistical decision theory (SDT) proposed by Wald [28]. Statistical Decision Theory (SDT) is a framework used in statistics and decision analysis to make optimal decisions under uncertainty. It is based on the idea of making decisions based on a set of probabilities and utilities, which quantify the likelihood of various outcomes and the value of those outcomes, respectively. The field of statistical decision theory deals with the process of decision-making in situations where statistical information (data) is available, providing insight into some of the uncertainties involved in the decision-making process [28].

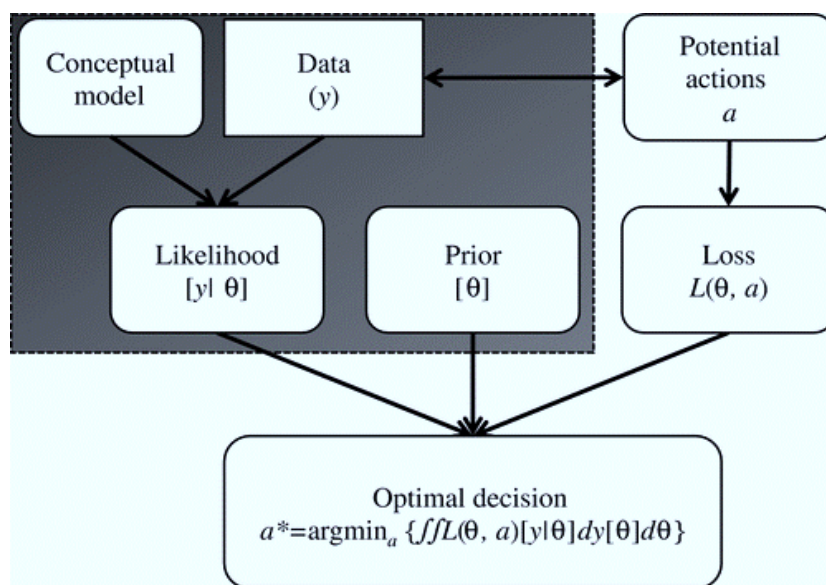
Wald proposed this theoretical approach as the foundation for statistical sequential analysis, which in turn resulted in the development of procedures in statistical quality control. The theory lingers on the idea that the accuracy of inferences made from statistical data could have impact on classical decisions. Although the theory and practice of Statistical Decision Theory (SDT) have been primarily advanced and published in areas such as mathematics, statistics, operations research, and other decision sciences, their incorporation in educational research has been limited [29,30].

Reading closely from the study of Bacci and Chiandotto [19] and Huang et al. [31] who also provided a clearer understanding of the theory, the impression and the understanding gathered is that statistical decision theory provides the theoretical background to approach decision theory from a statistical perspective. Therefore, any decision made from statistical perspective or analysis could have consequences on its users [19,30,31]. Further riding on the theory, the statistical decision theory that Wald formulated deals with the process of decision-making using sample data. Wald's idea of a statistical decision function (SDF) encompasses all types of mappings that arena the form of (data-decision). The theory preserves the assumption that how data are handled could influence the decision of the data [19].

Statistical decision theory (SDT) originated from the problem of a planner, decision-maker or agent who has to select an action that affects welfare depending on an unknown state of nature, as explained in theoretical analysis. The researcher identifies possible states and makes a decision without knowing the actual state. Wald enhanced this standard problem by introducing sample data that could provide information about the true state. He analysed the selection of a statistical decision function (SDF), which converts every potential data realisation into a feasible action. It is suggested that the assessment of SDFs as procedures chosen before the data observation, defining how the researcher will utilise any possible data.

Statistical Decision Theory (SDT) is a framework for making decisions based on statistical data. In a study like this, SDT can be applied to make decisions about hypotheses, models, and data analysis methods. For example, in a clinical trial comparing the efficacy of two treatments, SDT can be used to make decisions about whether the difference in outcomes between the treatments is statistically significant or not. The researcher can define a decision rule based on SDT that specifies the threshold for rejecting the null hypothesis of no difference between the treatments. This decision rule can take into account factors such as the sample size, effect size, and the acceptable level of risk of making a Type I error (rejecting the null hypothesis when it is actually true) [19].

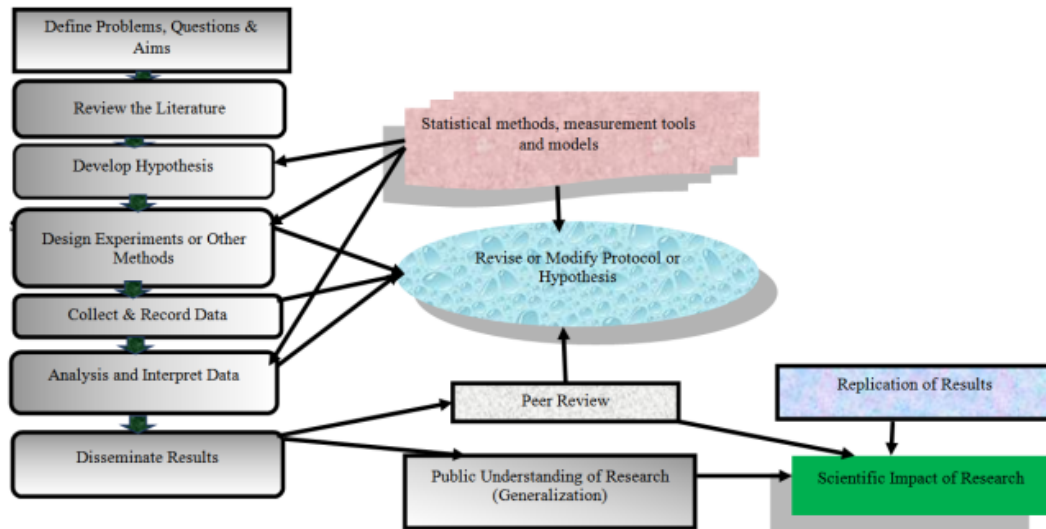
Statistical Decision Theory can also be used to select the most appropriate statistical model for a given dataset. The researcher can use decision theory principles to compare different models based on their goodness of fit, complexity, and predictive accuracy, and choose the model that best balances these factors. Overall, SDT provides a framework for making informed decisions based on statistical data, and can be applied in various stages of a research study to optimize the quality and validity of the results. The schematic diagram of statistical decision theory is depicted in Figure 1.



Source: Adopted from Wald (1950)

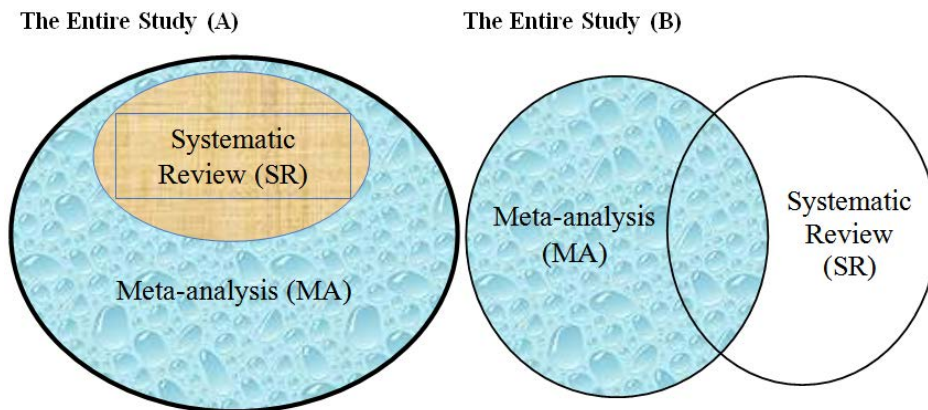
Figure 1. Schematic of Statistical Decision Theory (SDT)

Conceptual Framework



Source: Adapted from Gardenier and Resnik (2002)

Figure 2. Conceptual Framework (The Role of Statistics in the Scientific Method)



Source: Adapted from Giang et al. (2019)

Figure 3. Schematic of how Systematic and Meta-analysis were Apportioned in Study "A"

4. Methods

Research Approach

This study employed meta-analysis approach based on the assumptions that they are considered as forms of evidence-based practice methodological approaches. Evidence-based practice (EBP) is the process of integrating the best evidentiary information available with clinical expertise [32,33,34]. With this in mind, systematic review (SR) with meta-analysis (MA) as research approaches were employed to help combine all the available evidence fulfilling pre-determined criteria to answer the formulated research questions.

These approaches are recognised as being a crucial component of the practice of evidence-based research in order to obtain the highest level of evidence to formulate recommendations for theory and practice [32,33,34]. Systematic reviews and meta-analysis can provide a comprehensive and robust synthesis of the available evidence on a particular research question, allowing researchers and practitioners to make evidence-based decisions. However, they also have limitations, such as the

potential for publication bias and the heterogeneity of the included studies [32,33,34].

Meta-analysis is a statistical method that involves combining the results of similar quantitative studies, regardless of their level of significance. This method involves first calculating a standard effect size for each study and then pooling the effect sizes to generate a summary effect size. This approach has been highlighted by scholars such as Akhter et al. [33], and Li et al. [35]. According to Bolland et al. [36], "meta-analysis is created out of the need to extract useful information from the cryptic records of inferential data analyses in the abbreviated reports of research in journals and other printed sources." Prior to this understanding of meta-analysis, Smith and Glass (2000) plethora works on meta-analysis caught the attention of many social science researchers, and meta-analysis has gained respect across the social and medical sciences as a valid and rigorous methodological approach.

In employing these approaches, Giang et al. [37] each systematic review with meta-analysis should be designed and planned carefully. Both meta-analysis and systematic review are valuable research approaches that can provide

robust and evidence-based insights into a particular research question. They both require rigorous and transparent methods to ensure that the findings are valid, reliable, and relevant to the research question at hand. With this guidance and the propositions of Giang et al. [37] and Akhter et al. [33], Figure 3 depicts how meta-analysis and systematic review were blended and employed in the study. The study depicted the entire study “A”, where systematic review (SR) was only used as a subset of the entire study. This implies that the study was largely grounded in the meta-analytic approach.

5. Research Design

Under the umbrella of meta-analysis, the quantitative approach was used for the study. The rationale is based on the assumption of Zhang and Creswell [38] who asserted that quantitative approach in meta-analysis aims to examine a large amount of quantitative data where the researcher aimed at using statistical analysis of the data from independent primary studies focused on the similar question, which aims to generate a quantitative estimate of the studied phenomenon. The researcher employed the approaches based on the rationale that they help in synthesising data from multiple studies, meta-analysis provides a more precise estimate of the effect size, with increased statistical power to detect small but meaningful effects that may be missed in individual studies. This increased precision can help to inform clinical and policy decisions and guide future research. Moreover, the use of systematic and transparent methods to identify, appraise, and synthesize studies in a systematic review reduces the risk of bias and increases the credibility of the review. This makes it a valuable tool for decision-makers and researchers seeking to understand the evidence base for a particular intervention or phenomenon.

Population and Sampling Procedure

University of Cape Coast Institutional Repository (Data Location)

The data were extracted and managed from the UCCIR. The UCCIR is a digital service that collects, preserves, and distributes digital material. Repositories are important tools for preserving an organisation's legacy; they facilitate digital preservation and scholarly communication. An institutional repository (IR) is a digital collection of scholarly and research outputs created and managed by academic institutions, research organisations, or government agencies. IRs typically provide open access to a range of materials, including journal articles, conference papers, theses and dissertations, reports, and datasets, among others.

Data Extraction and Quality Assessment

To extract the data for the study, the eligible theses (PhD and MPhil) deposited at the UCCIR went through a standardised data extraction and quality assessment process. The data extraction form was refined during the extraction of the first thesis to ensure that the forms are comprehensive. The team (the researcher and the trained reviewers) extracted and screened descriptive characteristics of the sample from each quantitative study.

The extracted data from eligible studies were compiled using the guidelines of the 2020 PRISMA chat flow. To ascertain this, the researcher employed and coached reviewers and research assistants that assisted in the data extraction and assessment. In the process, each reviewer was tasked to assess and find out if the extracted theses were within the desired discipline context and contained statistical analysis. Again, the reviewers were tasked to check if the extracted theses were conducted, supervised and approved by the assigned supervisors under the guidelines of School of Postgraduate Studies and Research at the University of Cape Coast, Ghana. These quality assessment methods by collective reviewers helped in ensuring the reduction of bias by gaining some level of accuracy and transparency. After these rigorous processes of data extraction and quality assessment, a sample of 778 theses met the criteria. In essence, the results reported in this study are based on a final sample ($n=778$) conducted theses deposited at UCCIR. This sample is a comprehensive representation of the theses related to educational research.

It must be noted that for the main (substantive) analysis, only the quantitative studies were used. This is because, only the quantitative studies had the psychometric indicators that could be used in estimating the desire parameters for the formulated research questions. In essence, the sample for the analysis of 778 was reduced to 752 (all the extracted studies of 778 minus only the qualitative studies of 26, that is, $778-26=752$). Figure 4 presents the 2020 Preferred Reporting Items for Systematic Review and Meta-analysis statement (2020 PRISMA) flow chat of how the data are extracted from the journals database.

Data Collection Procedure

To efficiently and sufficiently obtain the data, the researcher employed and coached reviewers and research assistants that assisted in the data extraction and assessment. In the process, each reviewer was tasked to assess and find out if the extracted theses were within the desire discipline (educational field) context and contained statistical analysis. Again, the reviewers were tasked to check if the extracted theses were conducted, supervised and approved by the assigned supervisors under the guidelines of School of Graduate Studies and Research in the University of Cape Coast, Ghana.

In essence, after these rigorous processes of data extraction and quality assessment with the assistance, a sample of 778 theses met the set criteria for the study. This, therefore, suggest that the results reported in this study were based on a final sample ($n=778$) conducted theses deposited at UCCIR. This sample is a comprehensive representation of the theses that were related to educational research. It must be noted that for the main (substantive) analysis, only the quantitative studies were used for the statistical analysis. This is because, only the quantitative studies had the psychometric indicators that could be used in estimating the desire parameters for the formulated research questions. In essence, the sample for the analysis of 778 was reduced to 752 (all the extracted studies of 778 minus the qualitative studies of 26, that is, $778-26=752$).

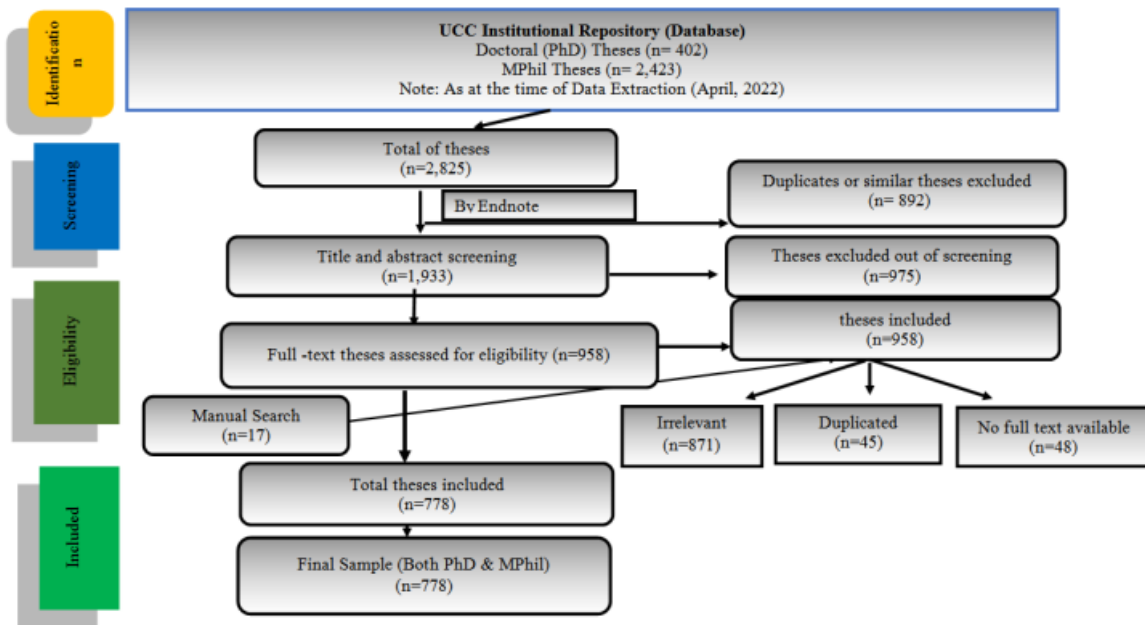


Figure 4. The 2020 PRISMA Flow Diagram on how the Data were extracted from the University Theses Repository

Data Processing and Analysis

In this study, the extracted data were analysed using several parameters through the help of some computer softwares such as STATA, CMA and SPSS (add-on-version of the version 26). During the analysis process in a meta-analysis, researchers should be knowledgeable about the nature of the data collected in each study, such as whether it is dichotomous or continuous, and select appropriate effect measures for comparison and estimation. Utilising computer software can facilitate the meta-analytic estimation process, including calculating the summary effect size, making corrections for potential biases, and identifying moderation effects.

Although spreadsheet programmes like MS Excel and statistical packages like SAS, AMOS, and LISREL are sometimes used for certain meta-analytic procedures, they are not optimal software solutions for meta-analysis studies because they lack important meta-analytic tools. For example, these programmes do not provide features such as forest plots and are unable to assign sample size-derived weights to individual effect sizes. It is better to use software that is specifically designed for meta-analysis, as recommended by Borenstein et al. [39], Iyengar and Greenhouse [40], and Aloe and Garside [41].

Subsequently, the deployment of specialised software such as Comprehensive Meta-Analysis (CMA), STATA, PLS-Smart and SPSS (add-on-version of the version 26) were used. These software (STATA, CMA and SPSS (add on version) were used based on the assumption that they could help in deriving parameters with the individual effect sizes [41,42,43]. In effect, these softwares assisted in estimating parameters such as Cochrane's Q-statistic, Orwin's fail-safe N and Cohen's d^a . They were also used to estimate parameters such as Tau-squared (T^2), I-squared (I^2), H-squared (H^2) and R-squared (R^2).

To analyse the guided research questions for the study, several parameters and models were employed with the help of the software. This was done to provide answers to the posed research questions. The first research question was on finding out the combined estimated

effect sizes of the extracted studies using Cohen's d^a and Hedge's g approach. The second research question focused on estimating heterogeneity analysis and weighted standard deviation for the studies where parameters such as Q-statistic, I-Squared (I^2), T-Squared (T^2), Pq, and Tau were estimated.

Table 1. How the Trend Data were Analysed

Indicators	Tools and Graphs Employed
Year Deposited at the UCCIR	Frequency and percentage and Error Bar (CI; 95%)
Thesis Level (PhD or MPhil)	Frequency and percentage and Error Bar (CI; 95%)
Type of Study Conducted	Error Graph (CI; 95%)
Validity Reporting Practices	Frequency and percentage and Error Bar (CI; 95%)
Type of Validity Evidence Reported	Frequency and percentage and Boxplot (CI; 95%)
Nature of the Instrument(s)	Error Bar (CI; 95%)
Reliability Reporting Practices	Error Bar (CI; 95%) and Boxplot (CI; 95%)
Type of Reliability Evidence Reported	Error Graph (CI; 95%) and Scatter Plot
Effect Sizes (ESs) (Coeff. Values)	Error Bar (CI; 95%) and Boxplot (CI; 95%)
Type of Analysis Employed	Frequency and percentage
Specified Confidence Intervals (CIs)	Error Bar (CI; 95%) and Boxplot (CI; 95%)
Assumption Testing Practices	Error Graph (CI; 95%) and Scatter Plot
Type of Assumption Tested	Error Bar (CI; 95%) and Boxplot (CI; 95%)
Sample Size (Observations)	Error Graph (CI; 95%) and Scatter Plot
Type of Design(s) Employed	Error Bar (CI; 95%) and Boxplot (CI; 95%)
Frequent Statistical Tools Employed	Error Graph (CI; 95%) and Scatter Plot

The analysis looked at the main or the substantive analysis based on the research questions. The analysis looked at the combined effect sizes (ESs), heterogeneity analysis and weighted standard deviation, the combined variance components, the random and fixed effect models and the meta-regression coefficients. In estimating theses, parameters such as p-values, z-values, Std Errors, UILL, CIUL and weights in percentages, Q-statistic, I-Squared

(I^2), T-Squared (T^2), Pq, Tau and WSD, Variances and Std Error, UILL, CIUL and df, FEM and REM, UILL, CIUL and df, B, SE, CILL, CIUL, Beta, Z and p-values were the focus. All these parameters were estimated at 95% level of confidence interval (CI;95%). Models such as Cohen's d and Hedge's g , Random and Fixed effect models and random effect model and method of moments estimation approach were employed. Table 1 depicts how the main data were processed and analysed using parameters and specified models.

6. Results

Assumptions Report-Test of Homogeneity

Before the substantive analysis for the study, this aspect of the analysis reported an assumption on the homogeneity of the data. In the assumption analysis, Cochran's Q statistic, Orwin's fail-safe N and Cohen's d^a were used to test for the homogeneity of the data. This was established the homogeneity of the data and pave way for the substantive analysis. In this, the researcher was interested in finding out whether the effect size measures in these studies are comparable. Specifically, they need to have the same or similar scale across extracted studies from UCCIR. Also, the assumption testing was necessary to assert that all the extracted studies are methodologically sound, i.e., data have been collected from a complete probability sample of a defined population, measurement has been valid and reliable, and the statistical analysis has been adequate. The accrued results on the assumptions are presented in Table 2.

Table 2. Test of Homogeneity

Parameters	Ch-Square (Q-statistic)	Df	Sig. (2-tailed)
Cochrane's Q statistic	11.334	3	.023**
Orwin's fail-safe N	12.891	3	.002**
Cohen's d^a	17.23	3	.003**
Overall Effect	55.013	769	<.001**

Source: Archival Data Retrieved from UCCIR (2011-2021)

* $p < .10$ (CI_{90%}), ** $p < .05$ (CI_{95%}), *** $p < .01$ (CI_{99%})

Table 2 presents the test of heterogeneity of the extracted studies that sought to test the homogeneity of the studies. The heterogeneity of the studies was estimated using Cochran's Q statistic, Orwin's fail-safe N and Cohen's d^a Tau-squared. From Table 2, the following values were recorded for each of the test; Cochran's Q statistic ($Q = 11.334$, $df = 3$, $Sig. = .023^{**}$, CI; 95%), Orwin's fail-safe N ($N = 12.891$, $df = 3$, $Sig. = .002$; CI; 95%) and Cohen's d^a ($d^a = 17.23$, $df = 3$, $Sig. = .003$; CI; 95%) with overall effect (OF=55.013; $df = 773$; $Sig. = .001^{**}$). As a result of the estimated values, it can be asserted that there is a statistically significant homogeneity between studies.

Test of Residual Heterogeneity

In the context of meta-regression, the test of residual heterogeneity is a statistical test used in meta-analysis to assess whether the variation in the effect sizes of the included studies is due to random error or to true differences in the underlying populations. Residual

heterogeneity refers to the variability in the effect sizes that cannot be explained by the sources of heterogeneity that have been accounted for in the meta-analysis. The most commonly used test of residual heterogeneity is the Q statistic, which is calculated as the weighted sum of squared differences between each study's effect size and the overall effect size. The Q statistic follows a chi-squared distribution with $K-1$ degrees of freedom, where k is the number of studies included in the meta-analysis.

A significant Q statistic (i.e., $p < 0.05$) indicates the presence of residual heterogeneity that cannot be explained by chance alone, suggesting that the studies are not sufficiently similar to be combined in a meta-analysis. In this case, the use of a random-effects model, which assumes that the true effect size varies between studies, may be more appropriate than a fixed-effects model, which assumes that the true effect size is the same for all studies. Another measure of residual heterogeneity is the I^2 -statistic, which represents the proportion of total variation in effect sizes that is due to true differences between studies rather than chance. The I^2 values range from 0% to 100%, with higher values indicating greater residual heterogeneity. An I^2 value of 50% or higher is generally considered to indicate substantial heterogeneity. To test of residual heterogeneity of the data, parameters such as Tau-squared (T^2), I-squared (I^2), H-squared (H^2) and R-squared (R^2) were used employed. The residual heterogeneity was used to test whether the studies are not accounted by the moderators. The computed values are presented in Table 3.

Table 3. Test of Residual Heterogeneity

Parameters	Para. notations	df	Sig. values
Tau-squared	$T^2 = .008$	4	.000**
I-squared (%)	$I^2 = 69.2\%$	4	.012**
H-squared	$H^2 = 2.34$	4	.003**
R-squared (%)	$R^2 = 32.1\%$	4	.005**

Source: Archival Data Retrieved from UCCIR (2011-2021)

* $p < .10$ (CI_{90%}), ** $p < .05$ (CI_{95%}), *** $p < .01$ (CI_{99%}), $n = 752$

Table 3 shows a test of residual heterogeneity of the extracted studies. The residual heterogeneity of the studies estimated, Tau-squared, H-squared, I-squared and R-squared values. From the results, Tau-squared ($T^2 = .008$, $df = 4$, $Sig. = .000^{**}$), I-squared ($I^2 = 69.2\%$, $df = 4$, $Sig. = .012$; $n = 752$) and H-squared ($H^2 = 2.34$, $df = 4$, $Sig. = .003$; $n = 752$) and R-squared ($R^2 = 32.1\%$, $df = 4$, $Sig. = .005$; $n = 752$). As a result of the estimated values, it can be asserted that there is statistically significant heterogeneity between studies.

Main (Substantive) Analysis

The main analysis focused on gathering data for the research questions. The first research question was on finding out the combined estimated effect sizes of the extracted studies using Cohen's d^a and Hedge's g Approach. The second research question focused on estimating heterogeneity analysis and weighted standard deviation for the studies where Q-statistic, I-Squared (I^2), T-Squared (T^2), Pq, Tau were given priority. The third research question was to estimate the combined estimated variance components. The fourth research question estimated random and fixed effect models of the extracted.

The last research question focused on estimating meta-regression of the extracted studies using the random effect model and method of moments estimation approach.

For emphasis, it must be reiterated that for the main (substantive) analysis, only the quantitative extracted studies were used for the analysis. This is because, only the quantitative studies had the psychometric indicators that could be used in estimating the desired parameters for the guided research questions. In essence, the sample for the analysis of 778 was reduced to 752 (778-26=752).

Research Question One: What are the Combined Estimated Effect sizes of Postgraduate Theses Deposited at UCCIR using Cohen's d and Hedge's g Approach?

In meta-analysis, the effect size is an important statistic used to quantify the magnitude and direction of the effect across all studies included in the analysis. A larger effect size indicates a stronger relationship between the variables or a larger difference between the two groups, while a smaller effect size indicates a weaker relationship or a smaller difference. Effect size in meta-analytic method is estimated to tell how meaningful the relationship between variables or the difference between groups is. It indicates the practical significance of a research outcome.

The concept of effect size goes beyond the notion of statistical significance and provides a way to demonstrate the importance and meaningfulness of research findings. The calculation of effect size is a statistical method that helps determine the practical impact and relevance of research evidence. This approach considers the power of statistical analysis to provide meaningful insights, irrespective of the emphasis on statistical significance. Against this backdrop, the researcher was interested in ascertaining the practical significance of a research outcome deposited at UCCIR. To achieve this, research question one was framed to estimate the combined effect

sizes (ES) of the extracted studies. In the quest to interpret effect sizes in meta-analysis, there are several effect sizes indices that are used as proxy for comparisons. This may include the Cohen's d , Hedge's g , Steiger's ψ (ψ), Cohen's f^2 , η^2 (eta – squared), Odds ratio (OR), relative risk or risk ratio (RR), Cramer's V, Glass's etc.

However, the study employed Cohen's d and Hedge's g , based on the rationale that they are preferred in meta-analysis since they deal with scientific studies. In estimating these values (Cohen's d and Hedge's g) for measuring the effect size, parameters such as the maximum and minimum effect size, estimated, p-values (CI, 95%), z-values, Std Errors, UILL, CIUL and weights in percentages were all estimated serving as a complementary analysis. The results are presented in Table 4.

The results of the combined estimated effect sizes of the studies are presented in Table 4. The combined outcome of minimum effect size of the study is estimated as (Min. ES=.378; $p=.000$; Hg=1.72, $z=12.20$; Std Err=.238; $n=752$) and maximum effect as (Max. ES=.430; $p=.012$; Hg=.812, $z=14.12$, Std Err=.623). The overall resultant effect was significant (Overall=.591, $p=.000^{**}$). In estimating effect size in a meta-analytic study Cohen's d , a large or very large effect size (Large=0.8, 79%, Very large=1.0, 84%, Huge=1.5, 93%, Very huge=2.0, 97%) means that research studies have practical significance implication, while a small or medium effect size (Small=0.2, 58%, Medium= 0.5, 59%) indicates limited practical applications. In essence, using indicated values, the results imply that the overall effect size (ES=.591) falls at the medium range and as such the extracted studies may have limited practical applications. In meta-analytic studies, forest plots are useful graphical displays summarising results from a meta-analysis. The results in Table 4 are also depicted in the forest plot in Figure 5.

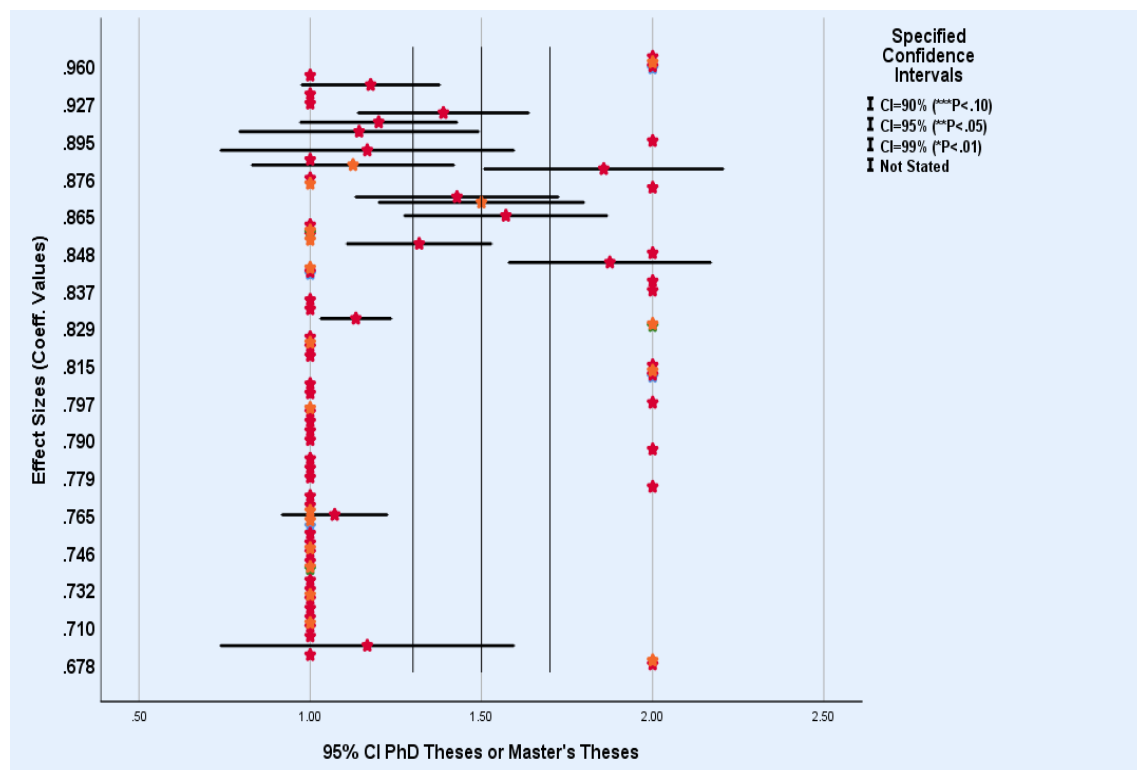


Figure 5. Forest Plot corresponding to the Figures in Combined Estimated Effect sizes (ES)

Table 4. Combined Estimated Effect Sizes (ESs)

Parameters	ES	p-value (CI, 95%)	z value	Hedge's g	Std Error	CILLCIUL	Weight (%)
Min. ES	.378	.000**	12.20	1.72	0.238	-0.43 1.27	11.13
Max. ES	.430	.012**	14.12	.812	0.623	-0.34 1.82	22.14
Estimated Overall Effect Size (Cohen's d^a = .591, $p > .000$ **)							

Source: Archival Data Retrieved from UCCIR (2011-2021)

Sample=752 (778-26)

* $p < .10$ ($CI_{90\%}$), ** $p < .05$ ($CI_{95\%}$), *** $p < .01$ ($CI_{99\%}$)

Key: Min. ES= Minimum Effect Size; Max. ES=Maximum Effect Size

Note: A large, very large, huge and very huge effect size means that research studies have practical significance, while a small or medium effect size indicates limited practical applications Using Cohen's d^a (Small=0.2, 58%, Medium= 0.5, 59%, Large=0.8, 79%, Very large=1.0, 84%, Huge=1.5, 93%, Very huge=2.0, 97%).

Research Question Two: What are the Combined Estimated Heterogeneity Parameters (Q-statistic, I^2 , T^2 , Pq, Tau) and Weighted Standard Deviation (WSD) of Postgraduate Theses Deposited at the UCCIR?

In meta-analytic studies, the estimated effects may vary across studies, partly because of random sampling error and partly because of heterogeneity. The fraction of variance that is due to heterogeneity is estimated by using heterogeneity parameters. It is against this backdrop that the aim of this research question was to estimate the heterogeneity analysis and weighted standard deviations of the extracted studies. For statistical analysis, heterogeneity parameters such as Q-statistic, I-Squared (I^2), T-Squared (T^2), Pq, Tau and weighted standard deviation (WSD) values were estimated using the Cohen's d^a and Hedge's g model. The result of the estimated heterogeneity parameters is presented in Table 5.

Table 5. Estimated Heterogeneity Analysis (Q-statistic, I^2 , T^2 , Pq, Tau) and WSD

Parameters	Estimated Values	p-value (CI, 95%)	df	z- value	Std Error	CILL	CIUL
Q-statistic	Q=7.768	.000	6	12.10	.344	-.13	2.21
I-Squared	I^2 =87%	.004	6	13.92	.234	-.12	1.23
T-Squared	T^2 =.627	.023	6	16.12	.245	-.23	.931
Pq	Pq=.934	.002	6	12.01	.812	.19	.813
Tau	Tau=.723	.004	6	30.01	.345	-.10	1.23
WSD	WSD=.075	.013	6	17.23	.201	-.08	2.28

Source: Archival Data Retrieved from UCCIR (2011-2021)
Sample**=752 (778-26)

Model**: Cohen's d^a and Hedge's g; * $p < .10$ ($CI_{90\%}$), ** $p < .05$ ($CI_{95\%}$), *** $p < .01$ ($CI_{99\%}$)

The estimated combined heterogeneity analysis (Q-statistic, I^2 , T^2 , Pq, Tau and WSD) of the studies are presented in Table 5. The Q-statistic was used to measure the variations of the extracted studies. The result of the Q-statistic is recorded as (Q=7.768, p =.000**, df =6, Z =12.10; CI =95%, n =752). A low p-value the Q-statistic indicates that there probably is some (undetermined) degree of heterogeneity. The I^2 was to measure and estimate the proportion of observed variance that reflects real differences in effect size of the extracted studies. In meta-analysis, if I^2 is low, then there is no heterogeneity to speak of and hence nothing to be explored in a subgroup or moderator analysis. The I^2 is recorded as (I^2 =87%,

p =.004**, df =6, z =13.92; CI =95%, n =752). If I^2 is large, then such an analysis is likely to be worthwhile. In this study, I^2 is 87%. This very high proportion suggests that the studies in this meta-analysis cannot be considered to be studies of the same population.

Again, both T^2 and Tau are measures of the dispersion of true effect sizes between studies in terms of the scale of the effect size. In this study, T^2 and Tau were used to estimate the variance of the true effect sizes. The results were estimated as T-squared; (T^2 =.627, p =.023**, df =6, Z =16.12; CI =95%, n =752), Pq; (Pq=.934, p =.002**, df =6, Z =12.01; CI =95%; n =752), Tau; (Tau=.723, p =.004**, df =6, Z =30.01; CI =95%; n =752), Weighted Standard Deviation; (WSD=.075, p =.013**, df =6, Z =17.23; CI =95%; n =752). The estimated values of the heterogeneity parameters suggest that there is probability some degree of heterogeneity in the extracted studies emphasis on the Q-statistics (Q=7.768, p =.000**, df =6, Z =12.10; CI =95%, n =752). These revelations may have accounted for the estimated medium effect size in the extracted studies (Cohen's d^a = .591, $p > .000$ **; Cohen's $d^a > 0.8$) analysed in research question one.

7. Discussion

Combined Estimated Effect Sizes of the Extracted Studies

In relation to the combined estimated effect sizes, the study found overall significant resultant overall effect size suggesting that the overall effect size is medium as such the studies may have limited practical applications. From the statistical analysis and results from the study, it was palpable that most of the estimated psychometrics reported non-significant effect sizes and this appears to have become more common place throughout the studies extracted for the analysis. This result suggests that the misplace of statistical analysis could lead to misplaced findings as such limiting practical and theoretical implications of studies.

Lending the findings to other related research evidences, it must be noted that in statistical analysis, data used for analysis and statistical assumptions should be given critical attention. In this statistical analysis of studies, the final decision on selecting statistical methods is made by following the algorithm for applying statistical methods to specific scientific fields. This algorithm allows for the quick and accurate selection of an appropriate method for statistical data processing [44]. In the ensued findings of this study, there were no other significant effects (with small effect sizes noted), and most of the other criteria appear to have improved as a result of the statistical inference report. These accrued results are not to imply that University of Cape Coast theses are in some way "missing the boat" with respect to statistical analyses, interpretation, inferences and documentations. However, most of the studies may have used fewer sample sizes and wrong statistical tools and this might have accounted for these limited practical applications of the extracted studies. Similarly, in the view of Chatterjee [45], it is recounted most statistical tests have long since been misinterpreted, significance

levels have been over-inflated, and type I error rates have been over inflated in a vast majority of studies.

It should be noted that recording significant effect size may add validity to many research studies, especially when there is a smaller sample size and effect sizes point to a much stronger outcome than p-values alone. These propositions of Hamrick et al. [46] and Frey [47] who posited that often, when employing a smaller sample, the effect size is a more accurate measure of the result of the experimental manipulation, which should sit well with both graduate students and researchers who have traditionally relied on the p-value as the lone determinant of the merit or significance of a study. As well, in the absence of effect size data, documentation of exact p-values facilitates post-hoc calculation of effect sizes [48,49,50] as required for meta-analysis. However, the study also acknowledges and recognizes that these statistical inferences on effect sizes may be battling a long history of trends in psychological research and publications.

In addition, many sophisticated statistical methods or models, such as analysis of covariances, mediation and moderation, repeated-measures analysis, logistic regression, ridge regression, lasso regression, polynomial regression, Bayesian linear regression, and survival analysis were seldom used or deployed in the extracted theses. An observation that was also made by several authors [35,51,52]. Making inferences from the study, it can be asserted that a large amount of data is not being efficiently analysed as such, much of the statistical information could be wasted and under reported. This mostly account for some misapplications and variations of some statistical misapplications and interpretation and assumptions in some studies [52]. From the data, it can be posited that UCC postgraduate students have made great efforts to employ statistical methods in their studies. However, the researcher cannot be overoptimistic because the situation appears satisfactory as the results showed a medium effect sizes indicating a limited practical and theoretical applications. This may not be limited to the Ghanaian context alone as some studies have statistical misapplications and interpretation could exist in some repositories and publication hubs of western countries [9,12].

Estimated Heterogeneity Analysis (Q , I^2 , T^2 , τ) and WSD of the Extracted Studies

The estimated combined heterogeneity analysis (Q -statistic, I^2 , T^2 , τ , P_q , τ and WSD) from the study found that most of the values of the parameters indicated that there is probability some degree of heterogeneity in the extracted studies. These results may have accounted for the small effect size in the extracted studies. This study showed that heterogeneity parameters could elucidate the variability in effect sizes that is not explained by the variance in sampling error. Meta-analysis examined in this research pointed out that there was a wide range in the amount and quality of information presented in meta-analysis for the studies taken from UCCIR. Making inferences on the ensued results from the study on heterogeneity analysis and weighted standard deviation, the impression created is that sometimes, the misapplication of statistics can also occur in the absence of erroneous or distorted students results. From the results

from the meta-analysis and systematic, it was found that some of the studies may have been driven by sampling errors that may have accounted for the variations and statistical misapplications.

Several authors [53,54] in their quest to track students' theses reached a conclusion that some of the studies may produce inconsistent and inaccurate results deal to the models they adopted. In these studies, the recommendation drawn is that to increase objectivity and establish more robust and accurate results, student researchers need to address such statistical issues such as excluding outliers, imputing data, editing data, cleaning data, or mining data. Again, Bahar et al. [55] asserted that the practices of inconsistent and inaccurate results are often practical, but it is important that researchers discuss their results honestly and openly to inform theory and practice. From the ensued findings from the study, it was gathered that there were some common types of misapplication of statistics in postgraduate theses and this could have accounted medium effect size of the extracted studies. Clearly, most the studies may have used or deployed improper statistical methods, techniques, or models in ways that produce distorted or artificial results. It was evident once more that the absence of disclosure of significant information regarding statistical methodology and assumptions was a systematic issue that could contribute to the production of inconsistent results. From a practical standpoint, the misuse of statistics could potentially breach various ethical obligations, such as the duty to uphold honesty, objectivity, and accuracy, and potentially even the duty to be transparent, as noted by Lucena et al. [56].

Drawing on the perspectives of other authors such as Bettany-Saltikov and Whittaker [18] and Worthy [57], it is important to note that statistical misuses that do not involve intentional deception could be attributed to honest mistakes, bias, incompetence, or serious deviations from accepted practice. In such cases, researchers who make frequent errors due to carelessness, lack of knowledge, or negligence may be considered to be lacking the necessary level of competence, whether in statistics or other domains, as noted by Bettany-Saltikov and Whittaker [18] and Worthy [57].

From this study, the researcher agrees with many authors [5,18] that have asserted that professionalism and integrity in education largely depends on sound statistical analysis and their interpretations. This is to assert that, scientific research in all disciplines requires a careful design and analysis of experiments and observations. Upshot to the above these authors have asserted that uncertainty and measurement errors are involved in most research studies. Clearly, the design, data quality management, analysis, and interpretation are all crucially dependent on statistical concepts and methods as such any abuse of statistical assumptions or procedures could affect some parameters leading to weak findings and conclusions [5,18]. For all of these reasons that have been accrued from the study and other related studies, it is important that researchers pay careful attention to the quality of statistical methods, methodological techniques, analysis performed, and report appropriately in their areas of jurisdiction and of competence. Good statistical work should be defended when it is attacked inappropriately.

Bad statistical work should be detected and corrected as appropriate [11,18].

The Outcome of the Tested Model

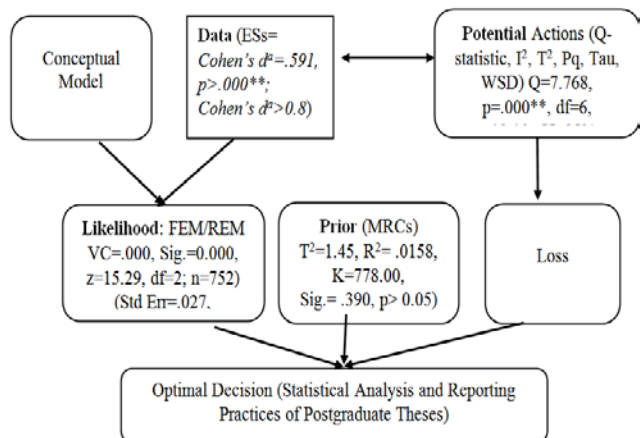


Figure 6. The Outcome of the Tested Model from Statistical Decision Theory (SDT)

8. Conclusions

Based on the findings of the study, the following conclusions are made from the study. Reasoning from the accrued results from the study, it is believed that good research depends on careful planning and execution. However, it can be concluded that there may be some statistical misapplications and interpretation that could be as a result of inadequate planning and execution of most postgraduate theses. This conclusion is reached on the medium effect size ensued from the extracted studies. Evidences from the study made the researcher to settle that there may be some wide variations in methods and reporting of the studies thereby casting doubt on the utility of the findings and recommendations of the studies.

From the findings of research question one, it can be concluded that there may be variability of the amount and quality of information reported in the extracted studies from UCCIR. This is to settle that some postgraduate theses may have rode on sampling error variances leading to different or weak conclusions and recommendations. Again, from research question two, it can be concluded that majority of the studies conducted and deposited at UCCIR are not far from producing similar results which could be due to similar approaches employed by the student researchers. This could be as a result of similar samples, similar respondents, similar studies and similar context.

Recommendations

While it is strongly perceived that there are still many important conceptual and theoretical issues relating to the application of statistics in research, it should be reiterated from this study and its discussions that more empirical research is required in the Ghanaian literature. In specific terms, from the accrued findings, the study put forth some recommendations for consideration: Reflection on the accrued results of the study, generally, there may be the need to improve upon the robustness and rigorousness of studies conducted under the guidelines of the School of Graduate Studies and Research (SGSR) at the

departmental, collegial and faculty level in the University of Cape Coast. This can be done when research mechanisms are put in place in tertiary institutions in Ghana to re-orient graduate students on emerging research issues specifically, statistical models, procedures and methodological assumptions. This will help position research in tertiary institutions in Ghana.

To expand literature in the Ghanaian context, it is recommended that the departmental, collegial and faculty level through the School of Graduate Studies and Research (SGSR) in the University of Cape Coast introduce meta-analysis and systematic review as a course. This could help expose students to some significant methodological and conceptual models in research. Internationally, it must be noted that most universities have embraced meta-meta-analysis and systematic reviews as fully enrolled courses for their postgraduate students. The rationale is to expose postgraduate students to powerful means of looking across datasets and providing scientific conclusions.

Data Availability

The primary data used to support the findings of this study are available from the corresponding author upon reasonable request and further perusal.

Declaration of Conflicts of Interest

No conflict of interest exists in this study. I wish to state categorically that there are no known conflicts of interest associated with this publication, and there has been no financial support for this work that could have predisposed the results of the study.

Ethics Approval and Consent to Participate

Ethics approval and consent to participate are not applicable in this publication as the study did not involve human subjects or data that require such approvals.

Funding

No funding was received for this study. The research was conducted independently without any external financial support.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my anonymous reviewers for their intellectual inspiration, stimulation, and constructive criticisms throughout the development of the manuscript. I also extend my profound appreciation to my research assistants for their assistance during data extraction, which contributed significantly to the completion of this study.

References

- [1] Baird, M. D., & Pane, J. F. (2019). Translating standardised effects of education programmes into more interpretable metrics. *Educational Researcher*, 4(8), 217-228.
- [2] Matthew, A. S. & Sunday, O. M. (2014). The Role of Statistical Software in Data Analysis. *International Journal of Applied Research and Studies*, 3(8), 1-15.
- [3] Petrişor, A. I. (2019). The use and misuse of statistics in research: Theory and practice. *Romanian Statistical Review*, 2(2), 59-70.
- [4] Žerovnik, J. (2015). About use and misuse of statistics in education: On mathematics exams at general matura in Slovenia. *Education Practice and Innovation*, 2(1), 1-7.
- [5] Khusainova, R. M., Shilova, Z. V., & Curteva, O. V. (2016). Selection of appropriate statistical methods for research results processing. *International Electronic Journal of Mathematics Education*, 11(1), 303-315.
- [6] Gardenier, J., & Resnik, D. (2002). The misuse of statistics: concepts, tools, and a research agenda. *Accountability in Research: Policies and Quality Assurance*, 9(2), 65-74.
- [7] Pigott, T. D. (2020). Missing data in meta-analysis. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed.). New York, NY: Russell Sage Foundation.
- [8] Calzon, B. (2021). *Misleading statistics examples discover the potential for misuse of statistics & data in the digital age*. News, Insights and Advice for Getting your Data in Shape.
- [9] Ntumi, S. (2021). Reporting and Interpreting Multivariate Analysis of Variance (MANOVA): Adopting the Best Practices in Educational Research. *Journal of Research in Educational Sciences*, 12(14), 48-57.
- [10] Sohn, S. Y., Rees, P., Wildridge, B., Kalk, N. J., & Carter, B. (2019). Prevalence of problematic smartphone usage and associated mental health outcomes amongst children and young people: A systematic review, meta-analysis and GRADE of the evidence. *BMC Psychiatry*, 19(1), 1-10.
- [11] Sabharwal, M. (2018). The use of soft computing technique of decision tree in selection of appropriate statistical test for hypothesis testing. In *Soft computing: Theories and applications* (pp. 161-169). Springer, Singapore.
- [12] Male, F., & Jensen, J. L. (2022). Three common statistical Lyu missteps we make in reservoir characterisation. *AAPG Bulletin*, 106(11), 2149-2161.
- [13] Mishra, P., Pandey, C. M., Singh, U., & Gupta, A. (2018). Scales of measurement and presentation of statistical data. *Annals of Cardiac Anaesthesia*, 21(4), 419-434.
- [14] Mandelbom, S., Manber, Z., Elroy-Stein, O., & Elkon, R. (2019). Recurrent functional misinterpretation of RNA-SEQ data caused by sample-specific gene length bias. *PLoS Biology*, 17(11), 370-381.
- [15] Ntumi, S., & Twum Antwi-Agyakwa, K. (2022). A Systematic Review of Reporting of Psychometric Properties in Educational Research. *Mediterranean Journal of Social & Behavioural Research*, 6(2), 53-59.
- [16] Bland, J. M. & Altman, D. G. (2017). Misleading statistics: Errors in textbooks, software and manuals. *International Journal of Epidemiology*, 17(2), 45-79.
- [17] Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486-489.
- [18] Bettany-Saltikov, J., & Whittaker, V. J. (2014). Selecting the most appropriate inferential statistical test for your quantitative research study. *Journal of Clinical Nursing*, 23(11-12), 1520-1531.
- [19] Matthew, H., Thiese, G., Zachary, G., Arnold, D., Skyler, S. & Walker, F. (2015). Statistics and ethics in medical research. Misuse of statistics is unethical. *British Medical Journal*, 19(2), 281-291.
- [20] Bacci, S., & Chiandotto, B. (2019). *Introduction to statistical decision theory: Utility theory and causal analysis*. Chapman and Hall/CRC.
- [21] Berger, J. O. (2013). *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media.
- [22] Tawfik, G. M., Dila, K. A. S., Mohamed, M. Y. F., Tam, D. N. H., Kien, N. D., Ahmed, A. M., & Huy, N. T. (2019). A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Tropical Medicine and Health*, 47(1), 1-9.
- [23] Kelley, J. M., Kraft-Todd, G., Schapira, L., Kossowsky, J., & Riess, H. (2014). The influence of the patient-clinician relationship on healthcare outcomes: A systematic review and meta-analysis of randomised controlled trials. *Plos One*, 9(4), 194-207.
- [24] Moreno-Peral, P., Conejo-Ceron, S., Rubio-Valera, M., Fernandez, A., Navas-Campaña, D., Rodriguez-Morejon, A., & Bellón, J. Á. (2017). Effectiveness of psychological and/or educational interventions in the prevention of anxiety: A systematic review, meta-analysis, and meta-regression. *JAMA Psychiatry*, 74(10), 1021-1029.
- [25] Fry, D., Fang, X., Elliott, S., Casey, T., Zheng, X., Li, J. & McCluskey, G. (2018). The relationships between violence in childhood and educational outcomes: A global systematic review and meta-analysis. *Child Abuse & Neglect*, 7(5), 16-28.
- [26] Garzón, J., Pavón, J., & Baldiris, S. (2019). Systematic review and meta-analysis of augmented reality in educational settings. *Virtual Reality*, 23(4), 447-459.
- [27] Rigabert, A., Motrico, E., Moreno-Peral, P., Resurrección, D. M., Conejo-Ceron, S., Cuijpers, P., & Bellón, J. Á. (2020). Effectiveness of online psychological and psychoeducational interventions to prevent depression: Systematic review and meta-analysis of randomised controlled trials. *Clinical Psychology Review*, 82(10), 19-31.
- [28] Ntumi, S. (2021). Reporting and Interpreting One-Way Analysis of Variance (ANOVA) Using a Data-Driven Example: A Practical Guide for Social Science Researchers. *Journal of Research in Educational Sciences*, 12(14), 38-47.
- [29] Wald, L. J. (1950). The theory of statistical decision. *Journal of the American Statistical Association*, 46(3), 55-67.
- [30] Williams, P. J., & Hooten, M. B. (2016). Combining statistical inference and decisions in ecology. *Ecological Applications*, 26(6), 1930-1942.
- [31] Insua, D. R., González-Ortega, J., Banks, D., & Ríos, J. (2018). Concept uncertainty in adversarial statistical decision theory. In *The mathematics of the uncertain* (pp. 527-542). Springer, Cham.
- [32] Huang, J., Yuan, Q., Zhang, B., Xu, K., Tankam, P., Clarkson, E., & Rolland, J. P. (2014). Measurement of a multi-layered tear film phantom using optical coherence tomography and statistical decision theory. *Biomedical Optics Express*, 5(12), 4374-4386.
- [33] Sackett, C. A., Kiepinski, D., King, B. E., Langer, C., Meyer, V., Myatt, C. J., & Monroe, C. (2017). Experimental entanglement of four particles. *Nature*, 404(6775), 256-259.
- [34] Akhter, S., Pauyo, T., & Khan, M. (2019). What is the difference between a systematic review and a meta-analysis? *Basic Methods Handbook for Clinical Orthopaedic Research*, 7(8), 331-342.
- [35] Mengist, W., Soromessa, T., & Legese, G. (2020). Method for conducting systematic literature review and meta-analysis for environmental science research. *Methods*, 7(3), 100-109.
- [36] Masuri, M. G., Othman, N., Wicaksono, G., & Isa, K. A. M. (2021). Translation and Validation of the Indonesian Version of SaringSikap Assessment. *Environment-Behaviour Proceedings Journal*, 6(16), 283-289.
- [37] Bolland, M. J., Grey, A., & Avenell, A. (2018). Effects of vitamin D supplementation on musculoskeletal health: a systematic review, meta-analysis, and trial sequential analysis. *The Lancet Diabetes & Endocrinology*, 6(11), 847-858.
- [38] Giang, H. T. N., Ahmed, A. M., Fala, R. Y., Khatlab, M. M., Othman, M. H. A., Abdelrahman, S. A. M., & Huy, N. T. (2019). Methodological steps used by authors of systematic reviews and meta-analysis of clinical trials: A cross-sectional.
- [39] Zhang, W., & Creswell, J. (2013). The use of “mixing” procedure of mixed methods in health services research. *Medical Care*, 51(8), e51-e57.
- [40] Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. Chichester, England: Wiley.
- [41] Iyengar, S., & Greenhouse, J. B. (2019). Sensitivity analysis and diagnostics. *Handbook of Research Synthesis and Meta-Analysis*, 8(7), 417-433.
- [42] Aloe, A. M., & Garside, R. (2021). Types of methods research papers in the journal Campbell systematic reviews. *Campbell Systematic Reviews*, 17(2), 1-9.
- [43] Brinckmann, J., Grichnik, D., & Kapsa, D. (2010). Should entrepreneurs plan or just storm the castle? A meta-analysis on contextual factors impacting the business planning performance relationship in small firms. *Journal of Business Venturing*, 25(1), 24-40.
- [44] Dalton, B., Bartholdy, S., Robinson, L., Solmi, M., Ibrahim, M. A., Breen, G., & Himmerich, H. (2018). A meta-analysis of cytokine

- concentrations in eating disorders. *Journal of Psychiatric Research*, 10(3), 252-264.
- [45] Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, 10(2), 8-13.
- [46] Chatterjee, K. (2017). Statistical fallacies in orthopaedics research. *Indian Journal of Orthopaedics*, 41(3), 37-46.
- [47] Hamrick, L. R., Haney, A. M., Kelleher, B. L., & Lane, S. P. (2020). Using generalizability theory to evaluate the comparative reliability of developmental measures in neurogenetic syndrome and low-risk populations. *Journal of Neurodevelopmental Disorders*, 1(2), 1-15.
- [48] Frey, B. (2018). *The SAGE encyclopaedia of educational research, measurement, and evaluation* (Vols. 1-4). Thousand Oaks, CA: SAGE Publications, Inc.
- [49] Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, 10(10), 535-549.
- [50] Colquhoun, D. (2017). The reproducibility of research and the misinterpretation of p-values. *Royal Society Open Science*, 4(12), 171-185.
- [51] Benjamin, D. J., & Berger, J. O. (2019). Three recommendations for improving the use of p-values. *The American Statistician*, 73(1), 186-191.
- [52] Heo, M., Kim, N., & Faith, M. S. (2015). Statistical power as a function of Cronbach alpha of instrument questionnaire items. *BMC Medical Research Methodology*, 15(1), 1-9.
- [53] Shrestha, N. (2020). Detecting multicollinearity in regression analysis. *American Journal of Applied Mathematics and Statistics*, 8(2), 39-42.
- [54] Apfelbaum, E. P., Phillips, K. W., & Richeson, J. A. (2014). Rethinking the baseline in diversity research: Should we be explaining the effects of homogeneity?. *Perspectives on Psychological Science*, 9(3), 235-244.
- [55] Omrani, H., Shamsi, M., & Emrouznejad, A. (2022). Evaluating sustainable efficiency of decision-making units considering undesirable outputs: an application to airline using integrated multi-objective DEA-TOPSIS. *Environment, Development and Sustainability*, 1(3), 1-32.
- [56] Bahar, B., Pambuccian, S. E., Barkan, G. A., & Akdaş, Y. (2019). The use and misuse of statistical methods in cytopathology studies: review of 6 journals. *Laboratory Medicine*, 50(1), 8-15.
- [57] Lucena, C., Lopez, J. M., Pulgar, R., Abalos, C., & Valderrama, M. J. (2013). Potential errors and misuse of statistics in studies on leakage in endodontics. *International Endodontic Journal*, 46(4), 323-331.
- [58] Worthy, G. (2015). Statistical analysis and reporting: common errors found during peer review and how to avoid them. *Swiss Medical Weekly*, 145(5), 1-7.



© The Author(s) 2024. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).