Journal of Research in Technical Careers
May 2018, Vol. 2, No. 1.

Meta-Analytic Procedures for Career and Technical Education Postsecondary Researchers and Practitioners

Conrad Oh-Young, Howard R. D. Gordon, Xue Xing, John Filler
University of Nevada, Las Vegas

Meta-analytic studies are syntheses of literature in which researchers use statistical means to summarize the findings presented across primary studies. They are of great interest in the fields of medicine and social sciences with numerous examples published in peer-reviewed journals. However, it appears that career and technical education (CTE) researchers are either not performing these research syntheses, or are not publishing their findings. Thus, there are three purposes to this manuscript. The first is to present CTE researchers and practitioners with a rationale as to why meta-analyses should be performed. The second is to provide guidelines that researchers and practitioners use to perform their own meta-analyses. The third is to provide suggestions that researchers and practitioners can use to disseminate the results of their meta-analyses. The manuscript concludes with a listing of suggested areas for future research.

Keywords: meta-analysis, systematic review, research synthesis, career and technical education, vocational education

Introduction

There is considerable interest in meta-analysis in the social sciences (Cooper, 2009), including the field of education. For example, in a review of meta-analyses in the field of education, Ahn, Ames, and Myers (2012) identified a total of 56 articles published across six American Educational Research Association (AERA) journals between 2000 and 2010. A brief search in ERIC for peer-reviewed journal articles published in the intervening time span (2011 through June 2017) using the keyword “meta-analysis” returned 1,506 entries. Though this search also included non-AERA journals, the amount is significantly more than the amount identified by Ahn et al., and clearly highlights the importance of meta-analytic research in the field of education.

Unfortunately, the importance of meta-analysis has not significantly pervaded the field of career and technical education (CTE). Rojewski, Asunda, and Kim (2008) investigated trends in the CTE literature from 2002-2004 using the peer-reviewed journals *Career and Technical Education Research*, *Journal of Career and Technical Education*, and *Journal of Industrial Teacher Education*. Their findings revealed that only one meta-analysis article was published during that time span. Gordon (2007) and Gordon, McClain, Kim, and Maldonado (2010) performed searches for CTE related meta-analyses using the ERIC and Academic Search Premier databases along with other methods, and concluded that CTE researchers may not be performing meta-analyses.

In 2017, the authors performed their own search for CTE related meta-analyses by focusing on those manuscripts published in *Career and Technical Education Research* and the *Journal of Career and Technical Education* from 2011 through 2016. After a review of 117 manuscripts, not a single meta-analysis was found to have been published during that time span.

Purpose and Objective. The continuing lack of meta-analytic studies published in CTE-focused journals suggests that CTE researchers and practitioners may not be aware of the advantages of meta-analysis, how to perform them, or how to disseminate their findings. Thus, the purposes of this manuscript are to provide CTE researchers and practitioners with a rationale as to the importance of performing meta-analyses, guidelines for performing meta-analyses, and guidelines for disseminating meta-analytic findings.

The Importance of Meta-Analyses

Types of Reviews of the Literature. In their book *Introduction to Meta-Analysis*, Borenstein, Hedges, Higgins, and Rothstein (2009) describe three types of reviews of the literature: narrative reviews, systematic reviews, and meta-analyses. A narrative review is a non-transparent and subjective summary of the literature (Borenstein et al., 2009), because the processes related to searching for and retrieving the literature may not be documented or consistent (e.g., a review that focuses only on studies published by well-known researchers). Narrative reviews are perceived as subjective because they involve the author assigning weights to each study that impact the possibility of its inclusion in the review. For example, an individual who is writing a narrative review on the perceptions of guidance counselors on CTE programs may choose to synthesize only those findings related to his/her areas of research, and focus less, or even exclude, those studies that do not. Because of these elements, narrative reviews of the literature have been referred to as qualitative in nature (Lipsey & Wilson, 2001).

The second type of review is the systematic review. These types of reviews differ from narrative reviews in that they are performed systematically with the intention that presentation of the methods used will be as transparent as possible. For example, an individual who is conducting a systematic review on the learning and teaching styles of secondary level CTE teachers may follow a systematic methodology of searching for, retrieving, and then synthesizing the literature as it pertains to the topic. This requires listing all appropriate search criteria used, how the methodological qualities of the studies were evaluated, and how those studies were analyzed by listing all important characteristics of the studies in a table. While systematic reviews are inherently more rigorous and less subjective than narrative reviews, the process of synthesizing the results is not entirely objective in that reviewers are still constructing a thematic story to fit the literature base.

Meta-analysis, the third type of review, is similar to a systematic review in that studies are systematically searched for, retrieved, and analyzed. It is an expansion upon the systematic review in that the data published in primary studies under review are also statistically analyzed. As described by Glass (1976):

> Meta-analysis refers to the analysis of analyses. I use it to refer to the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings. It connotes a rigorous alternative to the casual, narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding research literature. (p. 3)

Advantages of Performing a Meta-Analysis. In narrative reviews of the literature, the methods used to select studies for review may not be standardized. For example, a study may be selected because: (a) it is important to the field; (b) a well-known researcher was involved; (c) the findings support the viewpoint that is being conveyed; (d) it was easy to retrieve; (e) a large number of research participants were involved; (f) the results are statistically significant; or (g) it was performed by the same individual writing the review. In contrast, a meta-analysis would involve an exhaustive search where the authors would attempt to include all studies that meet a set of objective, predetermined selection criteria. Furthermore, the meta-analytic researcher would have explicitly described search and retrieval procedures, making it possible for other researchers to perform a follow-up analysis by replicating the original search procedures, and then including the studies published in the intervening years.

Another advantage to conducting a meta-analysis is related to the number of research participants available. For example, due to practical limitations a researcher who publishes the findings of a study on the effects of a technology-infused CTE teacher training program on classroom effectiveness may only be able to gain access to 40 research participants. Assuming there are multiple studies published on this topic, a researcher who seeks to perform a meta-analysis on this topic would review and analyze the findings from this prior study, but also across other studies. This could result in the synthesizing of data across hundreds, or even thousands of participants, thereby providing the meta-analyst with a larger base for inference.

Yet a third advantage relates to the analysis and presentation of findings. Other than the non-standardized selection criteria, the relative importance of literature within a narrative review may vary from article to article. As Borenstein (2009) has noted, this may be because both the review and analysis are author-centric, and thus subject to author bias. Since a meta-analysis utilizes statistical means of standardizing and then analyzing the data, the results are quantitative in nature. These results can then be compared across studies. Studies that stand out could be revisited, grouped together and further analyzed, or removed from the analysis entirely.

A fourth advantage is that findings may inform future research. Consider the hypothetical example of the technology-infused CTE teacher training program on classroom effectiveness. If a meta-analysis about this topic revealed that the more technology-rich a training program is, the more effective it is at producing high quality CTE teachers who remain employed for a minimum of 10 years, then researchers would want to consider conducting research that seeks to expand on those findings. Possible avenues of exploration may include: (a) investigating if certain subgroups respond to the training program in different ways (e.g., do individuals who are veterans outperform those who are not); (b) investigating if graduates of the program who work with high needs populations are more or less effective than their counterparts; or (c) determining what the threshold is before diminishing results are received. In contrast, if a researcher performed a meta-analysis on this topic and found that there were only a few studies in this area of research, with none having been published within the last 20 years, the meta-analyst would have highlighted a gap in the literature that could be revisited by CTE professionals.

Guidelines and Methodological Framework

The content presented is based on recommendations from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, 2015) and The Cochrane Collaboration (2011), as well as from experts in the field. Additional information beyond these recommendations are provided in the listing of video resources at https://goo.gl/awdFVF.

PRISMA. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) was created to provide guidelines that authors can use when reporting the findings of systematic reviews and meta-analyses (PRISMA, 2015). One of the available resources on the PRISMA website is a checklist that researchers can use to assist with preparing a manuscript for publication. This checklist is also available in Moher, Liberati, Tetzlaff, Altman, the PRISMA Group (2009).

Cochrane Database of Systematic Reviews. The purpose of the Cochrane Database of Systematic Reviews is to provide a collection of high-quality systematic reviews and meta-analyses related to the areas of health care and policy (The Cochrane Collaboration, 2017). While many areas of research presented within the Cochrane Database may not be directly associated with the field of CTE, researchers will benefit from use of the database in two ways. First, CTE researchers and practitioners are able to review other researchers' work. Second, the database houses the *Cochrane Handbook for Systematic Reviews of Interventions*, which provides guidelines for performing meta-analytic studies and systematic reviews.

Research Question and Purpose. Similar to primary research studies in which investigators identify and define sets of variables to examine, research questions, and purpose statements, meta-analysts must also do the same. At the minimum, CTE researchers and practitioners should strive to describe the meta-analysis's PICOS: Population: the population that is being studied (e.g., novice high school CTE teachers); Intervention: the treatment/independent variable (e.g., technology-intensive teacher training program); Comparator (or Control): the comparative group (e.g., teachers trained under the traditional teacher training program); Outcome: the dependent variable (e.g., measures of classroom effectiveness); and Study design: types of studies targeted for analysis (e.g., quasi-experimental group design studies).

Selecting an Appropriate Effect Size. CTE researchers and practitioners will need to select an effect size before data are analyzed. A flowchart on selecting an appropriate effect size is provided at https://goo.gl/DyQBNr with respective effect size formulas at https://goo.gl/x4bXbW. When selecting an effect size for meta-analysis, it will be beneficial if CTE researchers and practitioners are cognizant of the types of research and data available in the literature. If enough data are not available, then researchers may need to consider either reformulating their study's research questions or performing their own initial investigations into the matter.

Select a Model. There are two basic meta-analysis models to choose from: a fixed-effect model or a random-effects model. Borenstein et al. (2009) define use of the fixed-effect model as where the assumption is made "that there is one true effect size which underlies all of the studies in the analysis, and that all differences in observed effects are due to sampling error" (p. 61). Use of the fixed-effect model brings statistical power but limits generalizability (Rosenthal, 1995). In contrast, Borenstein et al. define use of the random-effects model as where the assumption is made "that the true effect could vary from study to study" (p. 61). It is the model of choice when the goal is to generalize findings to a larger population (Borenstein et al., 2009). Use of the random-effects model provides improved generalizability but less statistical power (Rosenthal, 1995).

Due to the differences between the studies that are to be analyzed (e.g., sample, settings, intensity of interventions, etc.), the true effect sizes will vary across them. Furthermore, researchers may want to generalize their findings to a larger population. Thus, the random-effects model is the most appropriate to use particularly for educational settings (Ahn et al., 2012).

Has the Study Been Performed? At this stage it would be appropriate to heed Moher et al.'s (2009) suggestion to investigate if a proposed meta-analysis has already been performed. In addition to searching databases that include CTE-related literature such as ERIC, two additional resources can be used to address this task. The first is a query of the PROSPERO database, a searchable database that can be used to determine if a given study has already been performed, is currently being performed, or is appropriate to carry out. If a study is not listed, then Moher et al. (2009) suggest registering the study. Upon registration, users will be prompted to provide specific information related to their proposed study such as target databases (e.g., ERIC), target intervention, and target population. A second resource is the Cochrane Database of Systematic Reviews. As stated previously, while the Cochrane Database primarily archives systematic reviews and meta-analyses of the literature related to health care and policy, CTE researchers and practitioners may find reviews of the literature related to their topics of interest.

Search Methodology. Similar to how researchers in primary studies determine, define, and describe their population, sample, and screening procedures, meta-analytic researchers must also define how they will search for data to collect and analyze. However, in the case of meta-analyses, these data are located in prior studies. Moher et al. (2009) provide a diagram that can be

Table 1. Types and Examples of Inclusion Criteria

Criteria	Examples
Years	Studies published between the years of 1980-2017.
Publication type	Studies published in peer-reviewed journals. Studies presented/published in CTE related conference proceedings.
Research Design	True experimental, quasi-experimental, mixed methods
Sampling method	Random sampling, stratified sampling, convenience sampling, snowball sampling
Journal(s) manuscript published in	*Career and Technical Education Research* *Journal of Agricultural Education* *Journal of Career and Technical Education*
Language of publication	English, Turkish, German, Chinese, Japanese
File type of manuscript	PDF, HTML, Word document
Participant demographics	Highest degree obtained for CTE teachers. Age of CTE students. Years worked as administrators.
Criteria appropriate to measured outcome	GRE test scores. Pre and post evaluations of teaching performance. Between subject versus within-subject comparisons.
Quality of data	Data from primary analysis. Data not included in prior meta-analyses.

used to assist with the documentation of the search and subsequent screening and evaluation processes.

Inclusion criteria. Meta-analytic researchers should define a priori what qualities studies will need to meet in order to be included in the analysis (Moher et al., 2009). A listing of possible inclusion criteria are presented in Table 1.

Exclusion criteria. In addition to the inclusion criteria, researchers may also choose to define specific criteria that would cause studies to be excluded from the analysis. Examples include longitudinal studies, studies in which sufficient data are not provided, studies that incorporate different designs that yield data that may not be appropriate for effect size calculations (e.g., mixed methods studies or single case design data), and studies that are secondary analyses of data.

Databases. In the 56 meta-analyses reviewed by Ahn et al. (2012), it was reported that all utilized electronic means (e.g., ERIC, PsycINFO, etc.) to search for studies. With the goal of the literature search to exhaust the current knowledge base, increasing the number of databases results in a wider net cast when searching for potential studies. This also has the potential to reduce biases such as retrieving only those articles published in well-known journals. Examples of commonly used databases are ERIC, Scopus, Academic Search Premier, PubMed, and PsycINFO.

Performing the search. Once the search criteria have been defined, the next step is to perform the search. While performing the search, it is recommended that all searches are performed together over a short period of time. This is because databases are continually adding new resources to their archives, resulting in searches performed during one period of time possibly returning fewer results than if that same search was performed a week or two later. For purposes of documentation, replicability, and transparency, researchers should also strive to keep a running log of when each search was conducted for each database, who performed the search, the number of results returned, and the types of results that were returned.

Retrieving files. It is important that whenever possible, PDF versions of the original articles are retrieved. If PDF files of the original articles are unobtainable, consider retrieving electronic versions directly from the publisher's websites, or retrieving the original print versions. The reason for this is that while older articles may be available as HTML full-text versions, these versions may sometimes omit content found in the original publications. Errata and retractions should also be retrieved.

Additional search strategies. In addition to the electronic keyword searches, CTE researchers and practitioners should consider incorporating additional search procedures. For example, the majority of the studies reviewed in Ahn et al. (2012) reportedly used backward and/or forward search strategies. Backward searches occur when the listing of references in an article are reviewed in order to identify other potential resources. Forward searches occur when a listing of articles that have cited a given study are reviewed. Databases such as Scopus and other selected publishers' websites offer features and metrics that can aid researchers with conducting both types of searches.

Another search strategy is to retrieve all of the published works of researchers who are experts in that area of study. Again, many databases and publisher websites offer metrics that can assist with this. ORCID

Table 2. Example of Hypothetical Posttest Data Collected

Study	Control Group (*N*)	Control Group (*M*)	Control Group (*SD*)	Treatment Group (*N*)	Treatment Group (*M*)	Treatment Group (*SD*)
Study #1	100	75	5.00	100	78	5.00
Study #2	80	80	6.00	80	85	6.00
Study #3	120	77	7.00	120	82	7.00

Note. The data provided in Table 2 depicts data from a hypothetical investigation into the effects of a technology infused CTE teacher training programs on classroom effectiveness. Posttest data from two independent groups were collected from each fictitious study.

(https://orcid.org) and sites such as ResearchGate, Google Scholar, and Academia.edu, can also be helpful.

Reviewing Studies. *Screening and eligibility.* Using the predetermined inclusion and exclusion criteria, researchers then filter through the studies that were retrieved. Studies should meet all inclusion criteria, and not meet any exclusion criteria. This is a time-consuming process that may take months or even years. Researchers should keep accurate and descriptive notes particularly if working with a team of individuals. After initial screening, researchers should consider reading the articles.

Quality of research methodology. The quality of the final results obtained from a meta-analysis is dependent on the quality of data that goes into those calculations (Lipsey & Wilson, 2001). Borenstein et al. (2009) refer to this as “garbage in, garbage out” (p. 380), in that errors made by the authors of the studies in the analysis will also be reflected in the final results obtained from a meta-analysis. Thus, a meta-analysis that is well-executed and uses data from high quality studies is more likely to result in findings that are of high quality. In contrast, a meta-analysis that is well-executed but uses data from poor quality studies may result in findings that are not meaningful. At its worst, these findings have the potential to lead to false conclusions, negatively impacting policy and practice. As summarized in Lipsey and Wilson (2001):

> Meta-analysts thus must carefully observe and code the key features of the studies judged eligible for the meta-analysis that bear on the validity and credibility of their results. If a large proportion of the studies are seriously flawed, corresponding cautions should be placed on any interpretation of the results and the analysis should be handled, especially carefully. (p. 157)

To address this issue, researchers should evaluate the quality of each study before including it in the analysis. Hancock and Mueller (2010) and What Works Clearinghouse (2014), are useful resources that can aid in this process.

Data collection. Once studies have been screened for inclusion and evaluated for research quality, the next step is to identify and collect data for later analysis. This process is referred to as coding and involves the determination of the types of data that will be collected, identifying who will collect these data, determining how these data will be recorded, collecting the data, and then evaluating the reliability of the data that were collected.

What to code? The first step in coding is to determine all of the possible elements of a given study that data will be collected on. Lipsey and Wilson (2001) describe two categories of data: empirical data and the general characteristics of each study. Empirical data are the quantitative data presented in the primary study. For example, if the end goal of the meta-analysis is to calculate a Cohen’s *d* summary effect, then the empirical data required to calculate a Cohen’s *d* effect size for each study will need to be collected. Table 2 provides examples of hypothetical empirical data that could be used to calculate a summary effect.

The second category of data refers to the characteristics within each study. Examples of these characteristics include the research design, participant demographics (e.g., age, gender, ethnicity, disability, and other categorical values), geographic location in which the study was conducted, the year of publication, and the journal in which the study was published. For thorough listings of characteristics, review the coding manual provided in Appendix E of Lipsey and Wilson (2001) and Chapter 7.3 in The Cochrane Collaboration (2011).

When attempting to determine what characteristics to code, CTE researchers and practitioners should consider the following: year the study was published, journal that published the study, participant demographics, research design, and sampling method. These data will be helpful when focusing on disseminating the results of the analysis. When deciding upon additional characteristics, researchers should consider first reflecting on how these data could be used to analyze and interpret the data beyond the overall summary effect. For example, a researcher performing a meta-analysis on the effects of a technology-infused CTE teacher training program on classroom effectiveness may also choose to collect data on participant demographics and geographic locations. During the process of collecting data on these characteristics, the researcher may discover that a majority of the studies in the analysis were conducted in rural settings and that the majority of the participants were male students. Because these data were coded, the researcher has the option of breaking the data into subgroups by gender and/or geographic location and then rerunning the analysis using only those data in the respective subgroups (e.g., male teachers in urban settings or female teachers in rural settings). Such data can also be used to identify trends in the literature.

Who will code? Lipsey and Wilson (2001) recommend that individuals who code should, at a minimum, have knowledge of both social science research and the field in which the meta-analysis is being conducted. Unfortunately, this may limit the available pool of coders to doctoral students and other selected researchers in the field (Lipsey & Wilson, 2001). And even this may not be sufficient.

Training coders. The level of training will depend on the coder's familiarity with research methodology and the field of interest. Sessions should include at least a few practice trials until coders are comfortable with both what to code and how to code. Though current technologies allow sessions to be held online, sessions that are held in-person should not be overlooked.

How will coding occur? It is imperative that coded data are stored in a format that is easily accessible for later analysis. While it has been recommended that researchers utilize a database program to assist with coding (e.g., Microsoft Access or FileMaker Pro; Lipsey & Wilson, 2001), there are also a variety of other software programs that could be used. First, a spreadsheet program could be used. Spreadsheets store data in a format that is easily transferable to statistical programs such as SPSS, R, or Comprehensive Meta-Analysis. Spreadsheet programs are also more accessible than database programs by offering researchers the option to share access online without the need for installation of additional software. Once access is shared, coders could independently populate the spreadsheets with data while the lead researcher periodically checks on their progress.

A second viable option would be to construct an online form that coders could respond to. Once team members are done responding, data would be exported into spreadsheet format for later analysis. The advantage to utilizing an online form over a spreadsheet is that the form could be constructed to limit coders to only those responses that are applicable (i.e., single answer array). Moreover, some online survey generator programs provide the option of configuring question items such that if certain criteria are met, then applicable question items are either displayed or not displayed. Contrast this with use of a spreadsheet in which coders may be continually scrolling from column to column, or row to row, as they are entering data.

Coding. All studies should be coded by at least two individuals. For a meta-analysis of 10 or fewer studies, the lead researcher should work with at least one other individual to independently code all studies slated for analysis. For more than 10 studies, researchers should divide into subsets for each coder to review while the lead researcher reviews the entire batch. For example, a researcher working on a meta-analysis on 30 studies may want to consider randomly assigning each coder 50% (N = 15) of the studies in the entire batch while he/she reviews all 30.

Communication. Ongoing communication between the lead researcher and the coders is important (Lipsey & Wilson, 2001), particularly when reviewing a study that is difficult to interpret. A few such scenarios are an article that describes the findings of more than one study, a study in which there was one control group and more than one treatment group (e.g., participants received the treatment at different levels of intensity), or an article that describes the results of a study that took place over multiple school years (i.e., multiple pre-post results with attrition throughout). The lead researcher will need to decide how to code these studies and then communicate this to the coders. Lipsey and Wilson (2001) recommend holding periodic meetings to address communication issues. Meetings could be held weekly or bi-monthly where team members provide updates, seek clarification, and review studies that were more difficult to interpret.

Researchers should also decide what to do in the event that coders are unable to retrieve necessary information from a given study due to lack of data. If this occurs, there are three options: (1) to remove the study from the analysis; (2) to attempt to make inferences; or (3) to attempt to fill the gap by collecting more data (i.e., contacting the author of the study). Whichever procedures are decided upon, they should be applied consistently and described as thoroughly as possible when preparing the manuscript for publication (e.g., what inferences were made, how prior authors were contacted, how data were requested from prior authors, what were the procedures if the authors did not respond or were not available).

Coder consensus. It is not uncommon for coders to become confused with inconsistencies among studies, focus on the wrong elements, incorrectly interpret variables, and/or mistakenly enter unnecessary data. This is particularly an item of note when coding information from studies that utilize different descriptors. For example, for the demographic element of ethnicity, one study may combine individuals who are Asian and Pacific Islanders into one category, another may combine Pacific Islanders with Native Americans, a third may list all three separately, and a fourth may differentiate between ethnicity and culture (i.e., participants may be of Asian descent but identify with Pacific Islander culture). As a result, checking for reliability among coders should be ongoing as opposed to summative (Lipsey & Wilson, 2001). Furthermore, the need for open communication and ongoing meetings should not be understated.

The Analysis. The analysis involves the calculation of effect sizes for each study. The effect sizes will then be used to calculate the summary effect. Figure 1 depicts the results of an analysis using a random-effects model and the hypothetical data presented in Table 2.

A spreadsheet template to assist CTE researchers and practitioners with calculations is provided at https://goo.gl/SeST8J. To use the spreadsheet, researchers should first download it, then enter the sample sizes, means, and standard deviations of two

Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Study#1	0.600	0.145	0.021	0.317	0.883	4.150	0.000
Study#2	0.833	0.165	0.027	0.510	1.156	5.056	0.000
Study#3	0.714	0.133	0.018	0.453	0.975	5.364	0.000
	0.707	0.084	0.007	0.542	0.872	8.392	0.000

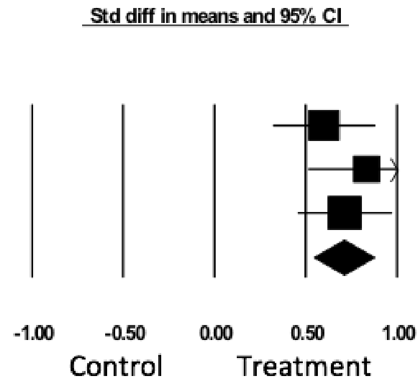

Figure 1. Forest plot diagram of the analysis of data from Table 2. The hypothetical results suggest that when analyzing data collected on 600 individuals across three studies, there was a significant difference between performance of participants in their respective groups favoring those who participated in the technology-infused CTE teacher training program with $d = 0.71$, $p < 0.001$, 95% CI [0.54, 0.87].

independent groups (the treatment and control groups) for each study. The spreadsheet will calculate effect sizes (*d*), the standard error, and the lower and upper CI limits for each study. The template will also calculate the summary effect.

Interpreting the Results. Researchers should reflect on the results as related to the original research question(s). The first step in doing so is to identify the summary effect calculations and the narrative interpretations as to what those data mean (see Table 3 for a listing).

Based on a meta-analysis of hypothetical data across three studies presented in Table 2, the technology-infused CTE teacher training program had a medium effect on classroom effectiveness favoring the treatment group: $d = 0.71$, $p < 0.001$, 95% CI [0.54, 0.87].

Not all calculations will yield effect sizes in the medium or large effect ranges. If this occurs, consider the following from Cohen (1988):

> In new areas of research inquiry, effect sizes are likely to be small (when they are not zero!). This is because the phenomena under study are typically not under good experimental or measurement control or both. When phenomena are studied which cannot be brought into the laboratory, the influence of uncontrollable extraneous variables (“noise”) makes the size of the effect small relative to these (makes the “signal” difficult to detect). (p. 25)

Researchers also want to consider that even if a small effect is received (e.g., $d = 0.2$), it still demonstrates an effect at one end of the continuum (Hedges & Hedberg, 2007). This is in comparison to a summary effect of $d = 0$, which suggests that there was no difference, or $d = -0.5$, which suggests a medium effect favoring the control groups.

Analyses of subgroups. If enough coded data are available, studies should be grouped into different subgroups based on categories that are of practical significance, and then further analyzed (Lipsey & Wilson, 2001; Schmidt & Hunter, 2015). Running meta-analyses on these different subgroups and then comparing and contrasting their effects may yield findings that are more insightful and more meaningful than the overall results (Lipsey & Wilson, 2001; Schmidt & Hunter, 2015). When considering the hypothetical meta-analysis on the effectiveness of a technology intensive CTE teacher training program, results could be subgrouped by the different factors that comprised classroom effectiveness and then further analyzed. For example, perhaps one subgroup of studies measured classroom effectiveness using performance on state assessments. These could be grouped together and analyzed.

Dissemination: Writing the Manuscript

With the meta-analysis complete, the next step is to disseminate the results. The following guidelines incorporate recommendations by Beretvas (2010), Moher et al. (2009), and The Cochrane Collaboration (2011).

During the writing of the manuscript, limitations related to the number of words/pages set by journals will be a constant issue. Researchers will need to cognizant of the balance between providing sufficient content to describe their findings and how those results impact the field, while at the same time presenting the methodology in a manner that is both transparent and replicable.

Title. A manuscript that disseminates the findings of a meta-analysis should have the word “meta-analysis” somewhere in the title. For example: A meta-analysis of

Table 3. Interpretation of Effect Size Values[a]

Effect size measure	Small effect size	Medium effect size	Large effect size	Very large effect size
Odds ratio	1.5	2.5	4	10
Cohen's *d* (or one of its variants)	0.20	0.50	0.80	1.30
R	0.10	0.30	0.50	0.70
Cohen's *f*	0.10	0.25	0.40	—
Eta-squared	0.01	0.06	0.14	—

Note. Adapted from "The Other Half of the Story: Effect Size Analysis in Quantitative Research," by J. M. Maher, J. C. Markey, D. Ebert-May, 2013, *CBE Life Sciences Education*, *12*(3), pp. 345-351. Copyright 2013 by The American Society for Cell Biology.
[a]Cohen, 1992, 1988; Rosenthal, 1996.

the effects of a technology-infused CTE teacher training program on classroom effectiveness.

Introduction. This section provides readers with the rationale as to why there is a need for this particular meta-analysis (Moher et al., 2009). Possible reasons include a call to action or changes in the political landscape. Researchers will also want to discuss the research that has already been conducted as well as provide information related to the current meta-analysis' PICOS (Population, Intervention, Comparator, Outcome, and Study Design; Moher et al., 2009).

Methods. *Registration information.* If the study was registered, provide the appropriate protocol registration information. Moher et al. (2009) suggest providing it in the methods section.

Model. A brief description of the meta-analysis model should be provided. This should include a definition of the model and a short rationale as to why it was selected.

Search procedures. Clearly and concisely describe how the literature was retrieved. Include the databases searched, years searched, keywords used, if specific journals were targeted, all inclusion and exclusion criteria, when the search was performed, and who performed the search. Moher et al. (2009) provide a flowchart that can be used to assist with describing these procedures. An example of a completed flowchart that describes the review of the CTE literature discussed at the beginning of this manuscript is available at https://goo.gl/Awi2Dg.

Coded data. Meta-analysts will be expected to provide a table that lists each study in the analysis. Descriptive information such as participant demographics, research design, sampling method, and types of outcome measures will need to be presented within the table. Empirical data from each study such as group sample sizes, effect sizes, standard errors, and confidence intervals should also be included. Space permitting, researchers may consider including additional coded data such as heterogeneity (see Rosenthal, 1995) or the results from the evaluation of study quality.

Inter-rater reliability. Though it has been noted that not all syntheses of the literature report inter-rater reliability (see Orwin & Vevea, 2009), Beretvas (2010) recommends that researchers should report the mean or median percentage agreement between coders. Other options such as kappa, weighted kappa, and intraclass correlation are also available (see Hoyt, 2010; Orwin & Vevea, 2009).

Results. The results will need to be presented in both numerical and graph formats. At the minimum, presentation of numerical results involves the summary effect, p-value, and confidence interval estimates (e.g., $d = 0.71$, $p < 0.001$, 95% CI [0.54, 0.87]). Though not discussed in this manuscript, researchers will also want to include results from other forms of statistical calculations where appropriate such as weighting (see Borenstein et al., 2009), power (see Borenstein et al., 2009), publication bias (see Schmidt & Hunter, 2015), or standard error.

The meta-analysis graph summarizes data gathered across multiple studies and presents it in a visual format. Researchers have the option of presenting graphed data in forest plot (Figure 1), funnel plot, box plot, or stem-and-leaf display formats. Researchers should include one graph for the overall summary effect and one per subgroup analysis. Programs such as SPSS, R, Comprehensive Meta-Analysis, or Microsoft Excel can be used to construct the graph.

Limitations. Meta-analytic research is not without limitations. With careful planning, methodological implementation of procedures, and persistent recordkeeping, the impact that these limitations have may be lessened. The following is a listing of limitations that should be addressed when preparing manuscripts for publication consideration.

1. Publication bias. Publication bias, referred to as the file drawer problem, references the notion that those studies that have positive effects tend to be accepted for publication more so than those with marginal or negative effects. If publication bias exists, then the results of those studies may influence the results of a meta-analysis.

Borenstein et al. (2009) acknowledges that while this is an issue, it is a problem that is reflected in all reviews of the literature, not just meta-analyses.

One strategy of addressing publication bias is to broaden literature searches to include studies that were not published in peer-reviewed journals (i.e., gray literature). Examples include conference presentations, reports (technical or grant-related), white papers, books, theses, and dissertations. Another strategy is to contact experts in the field to inquire about data from unpublished studies. A third is to statistically test for the existence of publication bias (e.g., trim-and-fill correction; see Schmidt & Hunter, 2015).

2. Apples and oranges. A second limitation is related to the differences between studies. For example, when considering three hypothetical studies that investigate the effects of an instructional intervention on the science achievement of high school students, one study may define science achievement as pre-post differences on a state assessment, a second may focus on student progress towards meeting Common Core Standards, and a third may focus on student GPA. How then could data from these studies be combined and analyzed as they would involve comparing an apple, an orange, and perhaps a pear? In response, Borenstein et al. (2009) makes the argument that while an apple, orange, and pear are being compared, all three are fruit. Therefore, the researchers are still investigating the effects of the instructional intervention, regardless if the fruit (how the construct of "science achievement" was measured), varied across studies. In practice however, while all three may be fruit, exercise in professional judgement should be used when performing the analysis and interpreting the results. To address this issue, researchers should consider performing an overall analysis, and then subgroup analyses on the same (or similar) measures of science achievement (e.g., a subgroup analysis solely on state assessment data).

3. Study quality. A third limitation is study quality (i.e., garbage in, garbage out). The presentation of the results from the evaluation of study quality should address this limitation. If a significant number of studies were identified as poor in quality, then this would be a contentious item for discussion that may also spur a call to action.

Discussion and Conclusion. This portion of the manuscript seeks to answer the question, "Where are we now that this meta-analysis has been conducted?" (Rosenthal, 1995, p. 190). Therefore, these sections should summarize and expand on the results while reflecting on how the findings from the meta-analysis impact the literature base, add to the field in general, and provide possible avenues for future research (i.e., Is there a disconnect between what the data suggests and what is currently being practiced?). For example, Moher et al. (2009) suggest providing discussion on how the findings impact the groups involved. Additional points of discussion include differences based on participant demographics, methodological consistencies or inconsistencies among studies, trends/gaps in the literature, or other moderating factors that could serve as rationale for future studies. Finally, researchers should look beyond statistical significance and provide discussion related to the practical implications that the results may have (Beretvas, 2010).

References. Barring publisher specific guidelines, APA style suggests that all studies included in the analysis should appear in the references section with an asterisk (*). Furthermore, the following statement appear below the references title, "References marked with an asterisk indicate studies included in the meta-analysis" (APA, 2010, p. 183).

Future Recommendations

CTE researchers and practitioners may want to consider engaging in meta-analytic research in the following subject areas as recommended by Gordon, Shaw, Xing, and Talib-Deen (2017): learning and teaching styles of CTE teachers, assessments of CTE programs, academic integration, and mentoring and preparedness of beginning CTE teachers. Focus could also be placed on analyzing data that may directly impact educational policy, for example, performing a meta-analysis on the effectiveness of instruction on student performance provided in CTE programs in charter schools versus CTE programs in traditional high schools. Within such a study, data could be coded, separated into subcategories, and then reanalyzed to focus on subsets of the overall population (e.g., virtual charter schools versus charter schools with in-person instruction).

Conclusion

Meta-analyses are statistical summaries of the literature that involve the synthesizing of empirical data collected across multiple studies. Having reviewed findings from prior articles (e.g., Gordon, 2007; Gordon et al., 2010) and performed a brief review of the literature, the need was deduced to present a rationale for the importance of meta-analyses, guidelines on how to perform them, and strategies on how to disseminate those findings. It is hoped that the contents of this manuscript provide CTE researchers and practitioners with information that can be used to perform meta-analyses that quantitatively synthesize where the field has been and illuminate future avenues of exploration within the field of CTE.

References

Ahn, S., Ames, A. J., & Myers, N. D. (2012). A review of meta-analyses in education: Methodological strengths and weaknesses. *Review of Educational*

Research, *82*(4), 436-476. https://doi.org/10.3102/0034654312458162
American Psychological Association (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, D.C.: American Psychological Association
Beretvas, S. N. (2010). Meta-Analysis. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 255-263). New York, NY: Routledge.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, NY: Wiley. https://doi.org/10.1002/9780470743386
The Cochrane Collaboration (2011). *Cochrane handbook for systematic reviews of interventions*. Retrieved from http://handbook.cochrane.org/front_page.htm
The Cochrane Collaboration (2017). *What is Cochrane evidence and how can it help you?* Retrieved from http://www.cochrane.org/what-is-cochrane-evidence
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159. https://doi.org/10.1037/0033-2909.112.1.155
Comprehensive Meta-Analysis (Version 2.2.064) [Computer software]. Englewood, NJ: Biostat.
Cooper, H. (2009). Hypotheses and problems in research syntheses. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis* (2nd ed.) (pp. 19-35). New York: Russell Sage Foundation.
Glass, G. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, *5*(10), 3-8. https://doi.org/10.3102/0013189X005010003
Gordon, H. R. D. (2007). *Meta-analysis research*: *A potential choice for CTE researchers and consumers*. Retrieved from https://eric.ed.gov/?id=ED498416
Gordon, H. R. D., McClain, C., Kim, Y., & Maldonado, C. (2010). Meta-analysis as a choice to improve research in career and technical education. *Career and Technical Education Research*, *35*(2), 103-116. https://doi.org/10.5328/cter35.208
Gordon, H. R. D., Shaw, S., Xing, X., & K. Talib-Deen. (2017, December). *Analysis of papers presented at career and technical education research and professional development conferences: 2009-2016*. Poster session presented at the Career and Technical Education Research and Professional Development Conference of the Association for Career and Technical Education Research. Nashville, TN.
Hancock, G. R., & Mueller, R. O. (2010). *The reviewer's guide to quantitative methods in the social sciences*. New York, NY: Routledge.
Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, *29*(1), 60-87. https://doi.org/10.3102/0162373707299706
Hoyt, W. T. (2010). Interrater reliability and agreement. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 141-154). New York, NY: Routledge.
Lipsey, M. W., & Wilson, D. V. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.
Maher, J. M., Markey, J. C., & Ebert-May, D. (2013). The Other Half of the Story: Effect Size Analysis in Quantitative Research. *CBE Life Sciences Education*, *12*(3), 345-351. https://doi.org/10.1187/cbe.13-04-0082
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), 1-6. https://doi.org/10.1371/journal.pmed.1000097
Orwin, R. G., & Vevea, J. L. (2009). Evaluating coding decisions. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis* (2nd ed.) (pp. 177-204). New York: Russell Sage Foundation.
Preferred Reporting Items for Systematic Reviews and Meta-Analyses (2015). *PRISMA transparent reporting of systematic reviews and meta-analyses*. Retrieved from http://www.prisma-statement.org/Default.aspx
Rojewski, J. W., Asunda, P., & Kim, S. J. (2008). Trends in career and technical education research. *Journal of Career and Technical Education*, *24*(2), 57-68.
Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, *118*(2), 183-192. https://doi.org/10.1037/0033-2909.118.2.183
Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research*, *21*(4), 37-59. https://doi.org/10.1300/J079v21n04_02
Schmidt, F. L. & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Los Angeles, CA: Sage Publications. https://doi.org/10.4135/9781483398105
What Works Clearinghouse (2014). Procedures and Standards Handbook Version 3.0. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf