ONLINE LEARNING
Volume 30, Issue SI2 (2026) · e5895
https://doi.org/10.24059/olj.5895

Managing Overlapping Review Studies in a Second-Order Meta-Analysis: Describing and Exploring Existing Strategies

Carl D. Westine and Chang Tong
University of North Carolina at Charlotte
Ting Sun
University of Utah
Tong Wu
Riverside InsightsFt

Abstract

This study explores different techniques for managing overlaps in second-order systematic reviews to highlight promising practices for open, distance, and digital education (ODDE) research. Second-order reviews, also known as overviews, provide unique evidence to better describe mean effects and assess their variability, but they also introduce new methodological challenges for quality control due to overlapping primary studies. Recent efforts to categorize and evaluate best practices for managing overlaps in other disciplines highlight significant inconsistencies in practice. As the frequency of ODDE systematic reviews and meta-analyses grows, there is a need to produce high-quality second-order studies to ensure policy makers can properly translate evidence into practice. Several leading overlap management methods are examined, including visualization tools such as citation matrices, bipartite citation networks, stacked bar charts, and intersection heatmaps; quantification tools such as corrected covered area (CCA), adjusted CCA and weighted CCA; decision-based and researcher-based approaches for resolving overlap. These approaches, largely developed in healthcare contexts, are illustrated using an existing dataset from a recent overview (Sun et al., 2025). CCA quantifies the degree of overlap among primary studies, while adjusted CCA further accounts for structural missingness arising when particular primary studies could not have been included in certain reviews because of differences in eligibility or timing. Although wCCA incorporates primary-study sample size into the assessment of overlap, it may be difficult to apply when information is unavailable. Existing thresholds for managing overlaps are largely based on practical guidance, lacking empirical evidence supporting specific cut-offs as decision thresholds.

Keywords: second-order meta-analysis, umbrella review, overlap, corrected covered area, citation matrix

Introduction

Research syntheses are a critical form of evidence used to shape research agendas. As this body of literature continues to expand, second-order reviews have emerged as an important approach for integrating findings across multiple systematic reviews. Various names have been used to describe this form of research: reviews of reviews, meta-syntheses, overview studies, umbrella-reviews or second-order meta-analyses (Cooper & Koenka, 2012; Polanin et al., 2017). Second-order reviews help researchers assess related problems with different interventions or examine studies with the same intervention but different outcomes (Becker & Oxman, 2008). While these reviews are not required to be purely quantitative, one form that is, a second-order meta-analysis, provides several distinct benefits like improving the estimation of true mean

Corresponding author: Carl Westine, cwestine@charlotte.edu
This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0).
© 2026 The Author(s)

effects and enabling assessment of the extent to which any true variance across the mean effect sizes exists (Schmidt & Oh, 2013).

Open, Distance, and Digital Education (ODDE) broadly encompasses teaching and learning processes in which educational technologies, digital media, and digital tools are used to deliver educational content and to support communication, interaction, collaboration, assessment, and evaluation (Zawacki-Richter & Jung, 2023). As research in this field expands across educational contexts, populations, delivery modes, and technologies, quantitative synthesis has become increasingly important for building cumulative evidence. For example, in distance education, Shachar (2008) noted that the abundant comparative literature remained inconclusive and often contradictory because studies varied in treatments, settings, measurement instruments, and research methods, and therefore advocated the use of meta-analysis. As meta-analytic evidence accumulates, returning to large bodies of primary studies becomes resource-intensive. Furthermore, within the fast-developing field of ODDE, interventions and contextual use can change quickly with new innovations and evolving uses across a variety of settings contributing to a literature base with imprecise definitions and descriptions of practice as well as contradictory evidence. This makes ODDE a natural context for growing second-order studies.

Second-order meta-analyses offer a more efficient means of synthesizing existing meta-analytic evidence and deriving cumulative conclusions (Young, 2017), as well as examining factors that lead to conflicting findings. Thus far, only a small but growing number of second-order meta-analyses have emerged across specific ODDE research areas (e.g., Sun et al., 2025; Borokhovski et al., 2022; Martin et al., 2022; Tamim et al., 2011; Yang & Xiang, 2024). However, as the volume and quality of systematic reviews and meta-analyses grows (Gnanasanjevi et al., 2025; Zawacki-Richter et al., 2025), more will be produced. Empirical trends documenting the emergence and importance of second-order reviews in other disciplines (Steenbergen-Hu et al., 2016) reinforce the important role that second-order studies will play in dictating practice within ODDE.

Although many second-order methods have been adapted from first-order studies, second-order studies do present unique challenges given the nested structure of the data across multiple levels. One unique challenge is overlapping primary studies, resulting in double counting of effects. In anticipation of a growing demand for second-order studies within ODDE, our efforts aim to describe and demonstrate several existing methods for identifying, examining, and removing overlaps. In doing this we hope to promote the efficient creation of high-quality second-order reviews and start a dialog to help researchers construct best practices within the ODDE context to ensure these comprehensive syntheses are most effective.

The Challenge of Overlaps

Given the nature of review studies as a form of synthesis, one challenge of any review study is addressing the influence of overlapping observations (Borenstein et al., 2009). Overlap occurs when the same participant effects appear in more than one included study. In a second order study, this can occur in a variety of ways and across more than one level (Polanin et al., 2017; Schmidt & Oh, 2013). This underscores the importance for researchers to clearly describe any identified overlaps and manage them in a transparent way.

Overlaps produce dependency, which in a statistical context can inflate Type 1 error and produce biased estimates. Dependency also manifests in several other related ways. For example,

some studies may report multiple outcomes on the same participants. Additionally, specialization can result in several similar studies being published by the same researcher. These other types of dependencies do not reflect the same double-counting of studies that occurs with overlap, and as a result can often be accounted for with statistical modeling. Various analysis techniques have been proposed and developed to model and account for dependency (e.g., robust variance meta-regression, and multivariate meta-regression, and multi-level modeling). These options offer valuable contributions to account for dependency when sample sizes are sufficiently large, but must be integrated with strategies to remove overlap stemming from double-counting.

Moreover, the amount of overlap that is present in a review study can be viewed as an evaluative criterion that has different levels of value depending on the study aims and researcher's epistemology. For example, as noted above, in a meta-analysis, where quantitative effects are weighted by study size, overlaps indicate double counting, a violation of statistical independence that is interpreted as undesirable. However, in qualitative reviews, overlaps may not be viewed as problematic because findings are not weighted by the study participants. Instead, overlaps may constitute an interesting study opportunity to characterize and explore.

We examine overlaps as they pertain to second-order meta-analyses where statistical independence is viewed as important and therefore minimizing the influence of overlap is paramount. We highlight and examine recent methodological advances from other fields to help streamline the process for ODDE researchers. By describing and demonstrating methods from this literature-base and reflecting on the process of our own experiences conducting second-order reviews we seek to provide greater transparency to both the practice and the root causes of the challenges associated with conceptualizing and resolving overlaps in second-order reviews within ODDE.

Overlaps in Second-Order Studies

As noted above, overlaps in second-order studies can occur in several ways. Just like in a meta-analysis, study participants may be the same in more than one primary study. This can occur in medical fields, where a specific diagnosis may be a driver of study inclusion and therefore increase the likelihood of the same participants appearing in multiple primary studies. Additionally, these patients may seek treatment for long periods of time and therefore be included in multiple research studies. Similarly, in education, overlaps might occur in the field of Special Education, where the presence of a specific (or uncommon) disability may induce greater likelihood of study inclusion. In ODDE, this type of overlap is less plausible. This is because in education the leveled and time-bound learning structure limits the opportunity and value of repeated interventions involving the same outcomes from the same participants. While cases certainly exist, they tend to be the exception. For example, researchers could conduct and publish multiple related (e.g., stacked treatment) studies that involve a variety of comparisons involving a common set of participants. However, in our own experience, these types of studies are uncommon.

More commonly, overlaps in second-order studies occur because there are primary studies that are included across multiple meta-analyses, and this should be expected in ODDE given the need for updating meta-analyses in response to innovations. This type of overlap can be wholly encompassing of study (such as what can occur longitudinally when one meta-analysis study successively builds off another) it can also vary depending on the inclusion criteria and search

strategies (e.g., one meta-analysis might include a few primary studies that are also included in only meta-analysis).

Theoretically, a single primary study might appear in one or up to all included meta-analyses. In a second-order meta-analysis because the primary studies are not the unit of analysis, they cannot easily be individually removed through deduplication. To reduce the amount of overlap to none or at least to an acceptably minimal level, whole meta-analyses must be discarded, potentially limiting the coverage of the second-order study.

Over the years, several strategies for managing overlapping primary studies have been employed. However, recent efforts to categorize practices for managing overlaps in other fields highlight significant reporting inconsistencies (Lunny et al., 2020; Pieper et al., 2014) and inaccuracies in the labeling of pairwise versus total dependencies (Bracchiglione et al., 2023). Even when overlaps are identified, techniques for resolving dependencies vary (Gutierrez-Arias et al., 2024), and remain largely untested. Additionally, nearly all strategies are based on reviews within the healthcare fields where managing overlaps necessarily differs given the size of reviews and the vast number that adhere to quality standards (e.g., Cochrane). The limited empirical evidence comparing the different strategies leaves researchers with unclear guidance (Lunny et al., 2021)

Different conceptualizations of overlaps can theoretically impact decisions on which studies to remove. Researchers must carefully examine and then communicate the nature of overlaps through visualization and calculation techniques to inform and defend any decisions on how best to resolve any issues with overlaps (Gutierrez-Arias et al., 2024).

To better position ODDE researchers to meet these needs, in the next section we will examine existing and emerging overlap management (quantification, visualization, and removal) approaches for second-order systematic reviews. We will also draw on our own experiences conducting second-order studies in ODDE to illustrate valuable process and provide methodological guidance to support the creation of second-order reviews in ODDE research.

Existing Examples of Managing Overlap from ODDE

Second-order meta-analyses in ODDE constitute an emerging research trend with only a few recent examples. We have identified only a few examples within the ODDE literature. As a result, limited guidance is available on how to handle overlaps in second order reviews within the field. To address this, we examine several approaches found in healthcare to the ODDE context. In this study, we describe and illustrate three categories of overlap management methods using Sun et al. (2025) as a running case. In the Sun et al. study, 15 relevant meta-analyses on multimedia were identified, but three studies did not provide enough details about the included primary studies to code for overlaps. Overlaps among the remaining 12 articles were examined, and one article was removed due to significant overlaps. By illustrating several leading methods for managing overlaps we aim to reflect on existing practices from healthcare and address the applicability of certain methods for use in ODDE. In doing so, we aim to help ODDE researchers more efficiently, effectively, and transparently manage overlaps in their second-order meta-analyses.

Table 1 highlights several examples and illustrates the various strategies for how overlap has been handled in existing second-order studies. Across these studies, approaches to primary-study

overlap varied considerably. Some studies reduced overlap through explicit selection rules, such as the 25% criterion (Demir et al., 2024; Tamim et al., 2011), whereas others combined overlap reduction with validation or robustness analyses to examine whether the overall findings changed after overlap was reduced (Yang & Xiang, 2024; Young, 2017). Other studies retained overlapping evidence in the main synthesis but conducted a de-duplicated validation analysis, chose not to formally quantify overlap based on prior validation evidence, or explicitly decided not to apply an overlap-reduction criterion (Bernard et al., 2014; Borokhovski et al., 2022; Hew et al., 2021). Ünal et al., (2026) quantified and visualized overlap by using GROOVE and CCA. Meanwhile, some studies considered statistical independence or acknowledged dependence as a limitation without formally quantifying primary-study overlap across meta-analyses (Sailer et al., 2024; Yi et al., 2025). Overall, these studies suggest that approaches to identifying, quantifying, and managing overlap remain highly study-specific, with no consistent strategy yet established across ODDE research.

Table 1

Second-Order Meta-Analytic Studies Identified in ODDE field.

Study	Context	Handling of overlap
Tamim et al. (2011)	Educational technology and learning	Assessed primary-study overlap and applied a ≤25% overlap criterion to select meta-analyses for inclusion.
Bernard et al. (2014)	Technology integration in higher education	Acknowledged primary-study overlap and noted Tamim et al.'s (2011) ≤25% overlap criterion but chose not to apply it, considering dependence from overlapping studies less severe in second-order meta-analysis.
Young (2017)	Technology-enhanced mathematics instruction	Assessed primary-study overlap and removed highly overlapping meta-analyses (>50%), reducing overall overlap from 20.8% to 13.8% and below the recommended 25% threshold (Cooper & Koenka's; 2012); a fully de-duplicated validation analysis produced a nearly identical effect (g = 0.38 vs. 0.39).
Hew et al. (2021)	Flipped classroom	Retained overlapping meta-analyses in the main second-order synthesis and conducted a de-duplicated primary-study-level analysis as validation.
Borokhovski et al. (2022)	Technology application in teaching and learning	Did not formally quantify primary-study overlap; excluded obvious duplicate/subset meta-analyses and relied on Tamim et al.'s (2011) validation indicating little added precision from further overlap assessment.
Demir et al. (2024)	Technology-based mathematics instruction	Retained overlap in the main synthesis and conducted a de-duplicated primary-study validation; the effect sizes were similar (g = 0.45 vs. 0.41), suggesting little impact of overlap on the overall estimate.
Sailer et al. (2024)	Technology-enhanced learning in higher education	Acknowledged overlap as a limitation, but did not formally quantify or adjust for primary-study overlap.

Study	Context	Handling of overlap
Yang & Xiang (2024)	Mobile learning and academic achievement	Assessed primary-study overlap and recursively removed the most overlapping meta-analyses to achieve <50% and <25% overlap thresholds (Sipe & Curlette, 1997; Tamim et al., 2011; Wilson & Lipsey, 2001); effect estimates remained similar across the full sample and reduced-overlap analyses (g = .652, .673, and .580).
Yi et al. (2025)	Technology-assisted foreign language learning	Considered data independence within individual meta-analyses but did not formally assess primary-study overlap across meta-analyses.
Ünal et al. (2026)	Artificial intelligence applications in education	Used GROOVE and CCA to quantify overlap; when overlap exceeded 10% (Bracchiglione et al., 2022), the narrower or less current meta-analysis was excluded.

Strategies for Managing (Quantifying, Visualizing, and Resolving) Overlaps

The potential problem of having overlaps typically arises early in the research process when screening for eligibility but is ultimately examined more fully as the final sample of studies to be used becomes more crystalized. To begin to address the overlaps, specific data preparation actions need to be taken to enable quantifying and visualizing the overlaps, before ultimately resolving the overlaps by removing select studies. As a first step, researchers must understand the nature of any overlaps that exist within their study. This requires deciding on an appropriate method for quantifying overlaps as well as describing through visualization, any overlaps that do exist.

Step 1: Understanding the Overlaps

Quantifying

Various strategies for quantifying overlaps exist. A recent systematic review of healthcare literature highlighted several prevailing approaches (Gutierrez-Arias et al., 2024). At a most basic level, researchers will sometimes calculate the percentage of overlapping primary studies among the systematic reviews. This can be reported for pairs of studies or for the total set of studies. Alternatively, a range of all pairwise overlaps may be reported among the studies included. Although conceptually easy to compute, one problem with simply reporting percentages is the amount of double-counting that can occur. For this reason, in the past decade, researchers have largely moved beyond simple percentages.

Calculating the Corrected Coverage Area. In 2014, Pieper et al. characterized the quantity of overlap using a corrected coverage area (CCA) to better account for primary studies that overlap with different sets of systematic reviews. The CCA removes the double counting of overlapping primary studies by considering each primary study once. While a bit more laborious to compute, this method is particularly relevant for comparing the amount of statistical independence in two different second-order studies. It provides a more consistent characterization of the aggregate amount of overlap that exists to allow for better comparison of overlaps between several second-order studies of different sizes. Recently, Ying et al. (2025)

presented a weighted CCA (wCCA), which expands the CCA calculation by weighting each systematic review by the sample sizes of the primary studies in each review. This gives the second-order researcher even more standardization in the amount of overlap and further improves comparability.

Missing Data. An important concern with quantifying overlaps is missing data. Missing data can occur in a variety of ways. For example, if a meta-analysis does not provide complete details on each of the primary studies included, then the extent of overlaps cannot be fully evaluated. This can be a major limitation when trying to assess statistical independence. For wCCA, the problem is compounded, as sample sizes for each individual primary study need to be known to properly weight each study.

Another type of missing data is structural missing data, for example due to the search ranges for systematic reviews not being the same. Primary studies in one systematic review may not be included in the timeframe of another study. For this reason, an adjusted CCA which removes these out-of-range studies from the calculation can be computed (Bougioukas et al, 2021).

Recent empirical research has pointed out the varying quality of systematic reviews in ODDE. According to Zawacki-Richter et al. (2025), one-third of articles in ODDE do not strictly follow PRISMA. Based on our own experience extracting information from ODDE studies for second-order reviews there is plenty of room for improvement in reporting the necessary information in primary studies. Complete information for assessing overlaps is not always provided, it can be inefficient to track down, or very difficult to extract consistently for each review due to a lack of enforcement of standardized reporting. With greater standardization in reporting of both primary studies and systematic reviews, the ODDE field will greatly enhance the assessment of the impact of overlaps in second-order studies.

Visualizing

Visualizing the overlaps between studies goes together with quantifying the amount of overlap. Multiple studies have introduced visual techniques to help describe overlaps that exist. Pieper et al. (2014) popularized the process of visualizing overlaps using a matrix. Gutierrez-Arias et al. (2024) found this was the most presented visualization. However, other techniques have been demonstrated as well. For example, Bougioukas et al. (2021) presented several other examples using traditional visualizations like Venn and Euler diagrams and network plots. Later, Bougioukas et al (2023) generated a tool to facilitate the efficient creation of a citation matrix and CCA heat maps, and Bracchiglione et al (2023) combined several visualization techniques to innovate how overlaps are presented using bar graphs.

Instructional Application Using Examples from ODDE - Resolving Overlaps

In this section we utilize data from a previous study by Sun et al. (2025) to demonstrate methods for understanding and describing the overlaps that are present in a study. We illustrate the process and reflect on several approaches for quantifying and visualizing the data.

The research team must first extract the necessary data from the articles. This means compiling a list of every primary study used in each systematic review to be included in the second-order systematic review. If there are overlaps, the same citations will be on the list two or

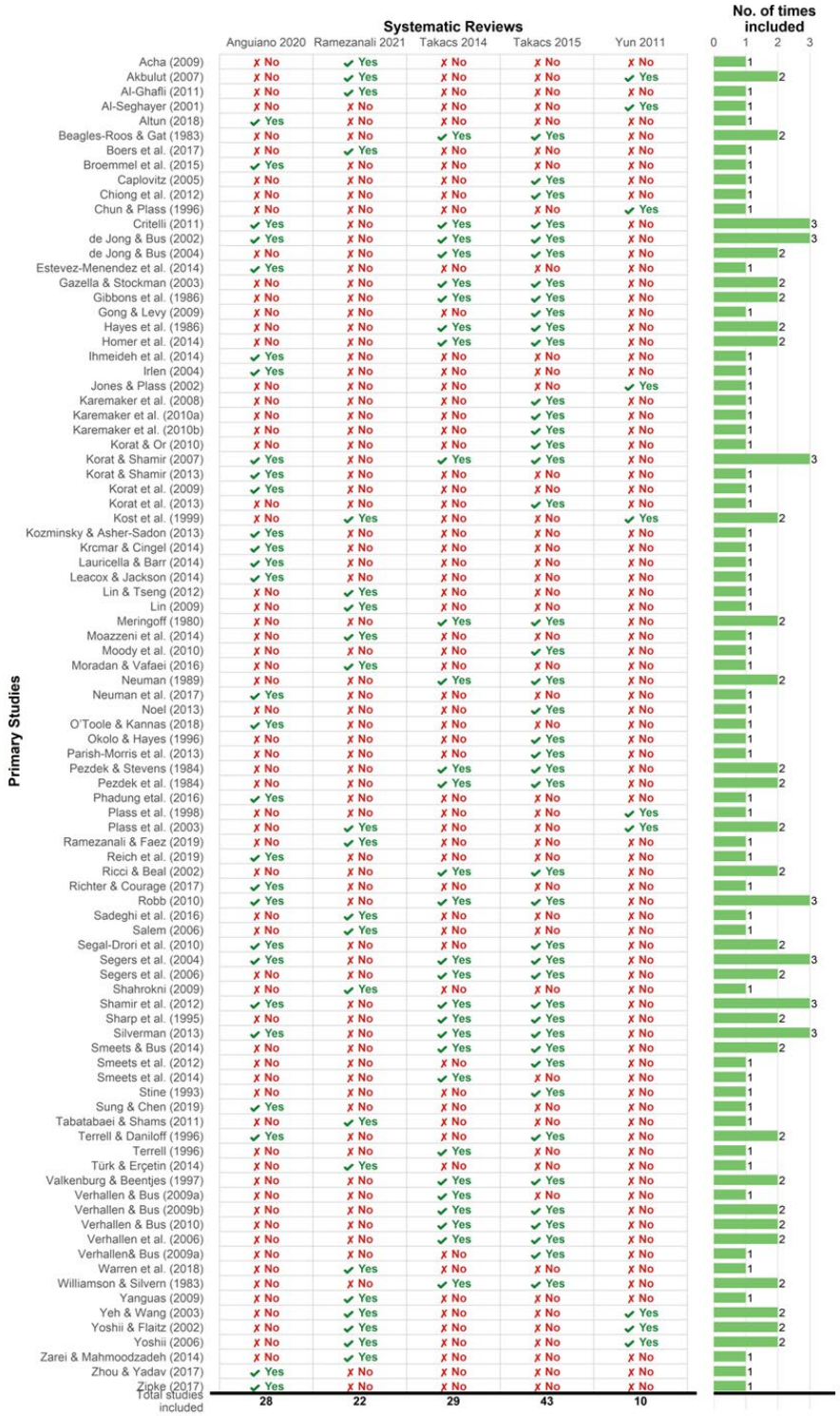
more times. Ideally all systematic reviews will include a full list of articles. If not, the researcher will need to proceed with only the subset of available articles, noting the limitation. For Sun et al. (2025), three studies did not have usable primary study lists, and needed to be removed from the overlaps analysis.

Citation Matrix Construction

The Citation Matrix is the foundation for calculating the CCA. It is a binary study-by-review table that documents the inclusion of primary studies across multiple systematic reviews. The leftmost column lists unique primary studies, and the subsequent columns represent individual systematic reviews and cells are coded to indicate inclusion or non-inclusion (See Figure 1). To create a citation matrix, researchers need to extract all primary studies from each review, identify unique studies after removing duplicates, and code their presence (1) or absence (0) across reviews. This structure enables identification and quantification of citation overlap across reviews.

Figure 1

Citation Matrix



Corrected Coverage Area

The Corrected Coverage Area (CCA) is a quantitative measure used to assess the degree of overlap among primary studies included in the systematic reviews (Pieper et al., 2014; Bougioukas et al., 2021). The CCA can be calculated using the following formula:

$$CCA = \frac{(N - r)}{(r * c - r)}$$

where *N* represents the total number of primary study occurrences across the systematic reviews (i.e., the sum of primary studies included within each review, including duplicates), *r* represents the number of unique primary studies included, and c indicates the number of systematic reviews. Pieper et al (2014) also provide guidance on interpreting the CCA percentages: 0-5% - slight, 6-10% - moderate, 11-15% - high, and >15% - very high. In Figure 1, *N* is identified as the sum of the values in the rightmost column (No. of times included, which has been augmented with visual bars for clarity), *r* is the number of individual study rows, and *c* is the number of study columns. For illustration purposes, Figure 1 only includes five of the 12 studies included in the overlap analysis. The CCA value can also be calculated using the ccaR() function in R once the full citation matrix has been prepared. Below, in Table 2, is the R output using the full set of data from Sun et al. (2025):

Table 2

R Output for CCA Calculation

reviews	N	r	c	CCA_Proportion	CCA_Percentage
12	326	285	12	0.01307815	1.3

This output indicates 326 total study occurrences (*N*) and 285 unique primary studies (r) across 12 reviews (c). The calculated CCA proportion is 0.013 (1.3%), indicating a very low level of overlap among reviews.

Adjusted Corrected Coverage Area

The adjusted CCA accounts for structural missingness, which occurs when primary studies were published after a given systematic review and therefore could not have been included (Bougioukas et al., 2021). In the overlap matrix, these structurally unavailable studies were coded as NA rather than as absent (i.e., coded as 0). This approach prevents underestimation of overlap due to temporal sequencing and provides a more accurate representation of redundancy across reviews (Bougioukas et al., 2021). The adjusted CCA is calculated as:

$$Adjusted\ CCA = \frac{(N - r)}{(r * c - r - X)}$$

where *N*, *r*, and *c* are the same as in the CCA calculation, and *X* is the total number of structural missingness in the citation matrix. The adjusted CCA can also be calculated in R using the ccaR() function after updating the citation matrix to account for structural missingness. Below, in Table 3, is the R output for the full set of data from Sun et al. (2025):

Table 3

R Output for Adjusted CCA Calculation

reviews	N	r	c	Structural_missingness	CCA_Proportion	CCA_Percentage
12	326	285	12	674	0.01665989	1.7

The ccaR() output for the adjusted CCA suggests that there are 674 structurally missing cells in the citation matrix. After accounting for the structural missingness, the adjusted CCA proportion is 0.017 (1.7%), which is slightly higher than unadjusted CCA. The results indicate that overlap across the reviews remains low even after adjusting for sparsity in the citation matrix.

Weighted Corrected Coverage Areas

The wCCA is a newly proposed extension of the traditional CCA, which incorporates sample-size–based weights (Ying et al., 2025). The wCCA can be calculated as:

$$wCCA = \frac{(N_w - r_w)}{(r_w * c - r_w)}$$

where N_w is the weighted sample size which consists of the sum of the square roots of the sample sizes of all primary studies, aggregated across all SRs and r_w represents the sum of the square roots of sample sizes for unique primary studies, where each primary study is counted only once. To manually compute the wCCA, the citation matrix needs to be appended to include an additional column which includes the sample size for each primary study listed in the first column. Compared with the CCA, wCCA provides a more precise assessment of the overlap by accounting for differential contributions across reviews. A similar extension to account for structural missingness could also be applied to create an adjusted wCCA.

Some practical challenges emerged when calculating the wCCA using ODDE articles from Sun et al. (2025). As Zawacki-Richter et al. (2025) pointed out, many systematic reviews lack transparent reporting, which can hinder the retrieval of key primary characteristics, including sample size. Pollock et al. (2017) also found similar data could be presented differently across reviews. Both were evident in our experience.

To illustrate the wCCA, we consider the same subset of five systematic reviews as above. However, only three systematic reviews reported the sample size for each primary study. Moreover, even when these sample sizes were reported, the identification of multiple related within-subjects comparisons required decision-making and interpretation of reported sample sizes to address the double-counting participants across conditions, which also affects statistical independence. Among the 87 primary studies included, this occurred in five articles. In addition, one primary study could not be retrieved.

To explore how different decisions might affect the wCCA, we applied three calculation strategies although differences in study focus and design could introduce bias. First, we used the sample sizes reported in the primary studies. Second, we relied on the sample sizes reported in the systematic reviews and supplemented missing values with those from primary studies. Third, we again used the review-reported sample sizes, but for missing cases we estimated values based

on the total reported sample and the average of the unknown portion. The resulting wCCA values for just these three studies were 0.1289, 0.1140, and 0.1146, respectively.

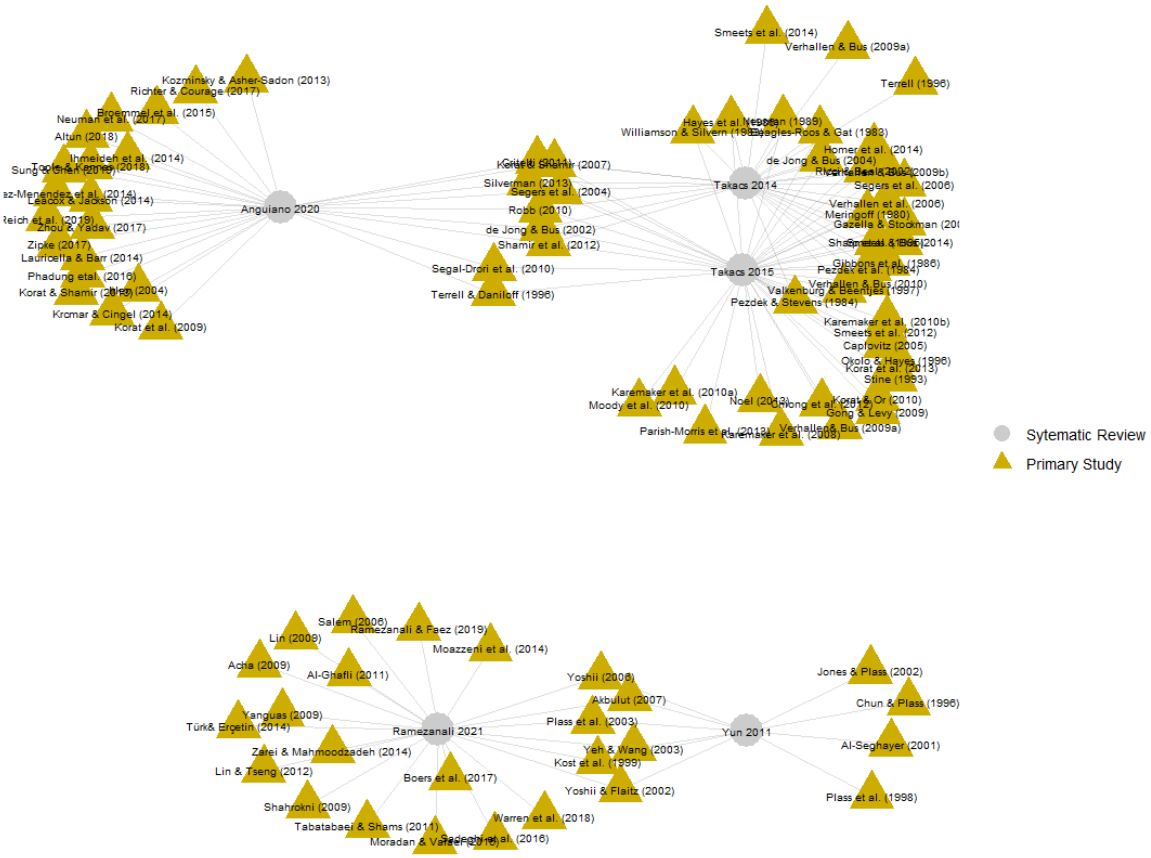
Supplemental Visualization Techniques to Gain Additional Clarity

While CCA-type values provide an overall summary measure, they do not indicate which specific reviews overlap most or where the overlap occurs. This can be gleaned from the citation matrix with careful inspection, though additional visualization strategies could be informative. Several visualization methods and illustrative examples were reviewed by Bougioukas et al. (2021). Figures and tables can be useful for depicting overlaps in terms of density and spectrum; however, most visualization techniques become difficult to interpret when the number of studies exceeds approximately five. For example, Venn diagrams, Euler diagrams, UpSet diagrams, and network-based approaches such as bipartite graphs have been used to illustrate study overlap. Although these methods can serve as helpful intermediate tools for exploring patterns of overlap, they are less suitable for reporting purposes because the visualizations often become overly crowded and difficult to interpret.

This limitation is illustrated in Figure 2, which presents a bipartite "cloud-like" citation network based on five first-order meta-analyses selected from Sun et al. (2025) as a demonstration. In the figure, grey nodes represent first-order meta-analyses, and yellow nodes represent primary studies. The visualization shows that Takacs (2014) and Takacs (2015) share a substantial number of primary studies, which is unsurprising given that they were authored by the same research team. These two meta-analyses also share numerous primary studies with Anguiano (2020). A separate cluster on the right side of the figure indicates that Ramezanili (2021) and Yun (2011) share a smaller number of primary studies. However, even with only five meta-analyses included, the primary study nodes appear densely clustered, limiting the clarity and interpretability of the visualization.

Figure 2

Bipartite citation network



Stacked Bar Chart

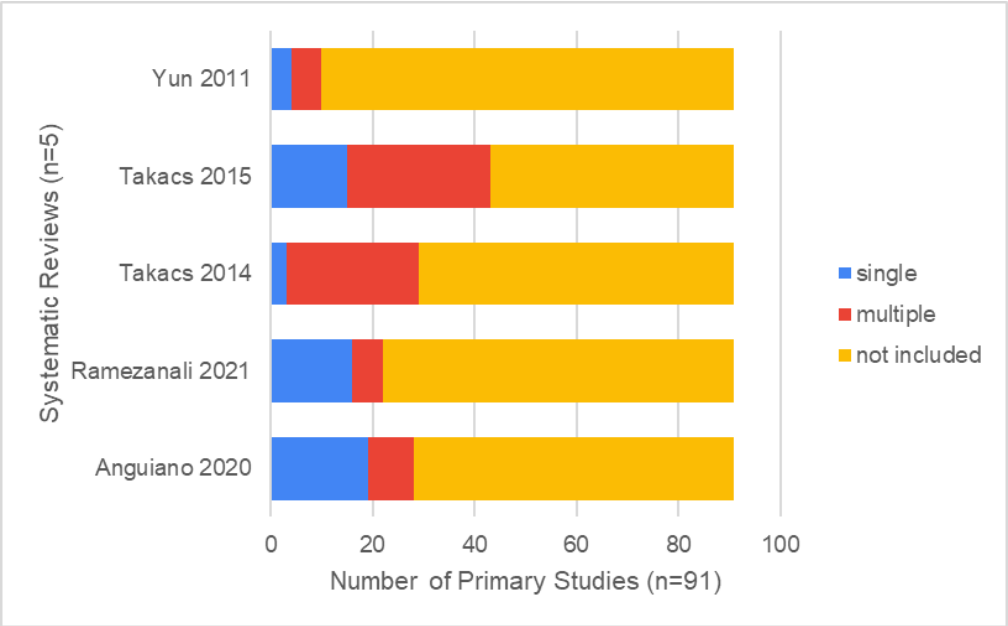
It is not uncommon for second-order studies to include many first-order meta-analyses, each of which may synthesize multiple primary studies (Lunny et al., 2020). Contrasted with the citation matrix (shown in Figure 1), which summarizes the number of times each primary study appears in the systematic reviews, it is necessary to also convey the percentage and nature of overlaps occurring within each systematic review. In this context, stacked bar plots are a useful visualization method for providing a relatively straightforward, high-level summary of overlap patterns, particularly in terms of the frequency of shared versus unique primary studies. However, bar plots cannot represent the specific structure of overlap (i.e., which primary studies are shared by first-order meta-analyses).

As shown in Figure 3 using the subset of articles from Sun et al. (2025), we labeled each bar segment using three categories, single, multiple, and not included, following Bougioukas et al. (2021). The blue segments (single) indicate primary studies that are exclusively included in each first-order meta-analysis. The red segments (multiple) represent primary studies that are included in the corresponding first-order meta-analysis as well as in one or more other first-order meta-

analyses. The yellow segments (not included) indicate primary studies that are not included in the corresponding first-order meta-analysis.

Figure 3

Stacked Bar Chart of Primary Study Inclusion Status



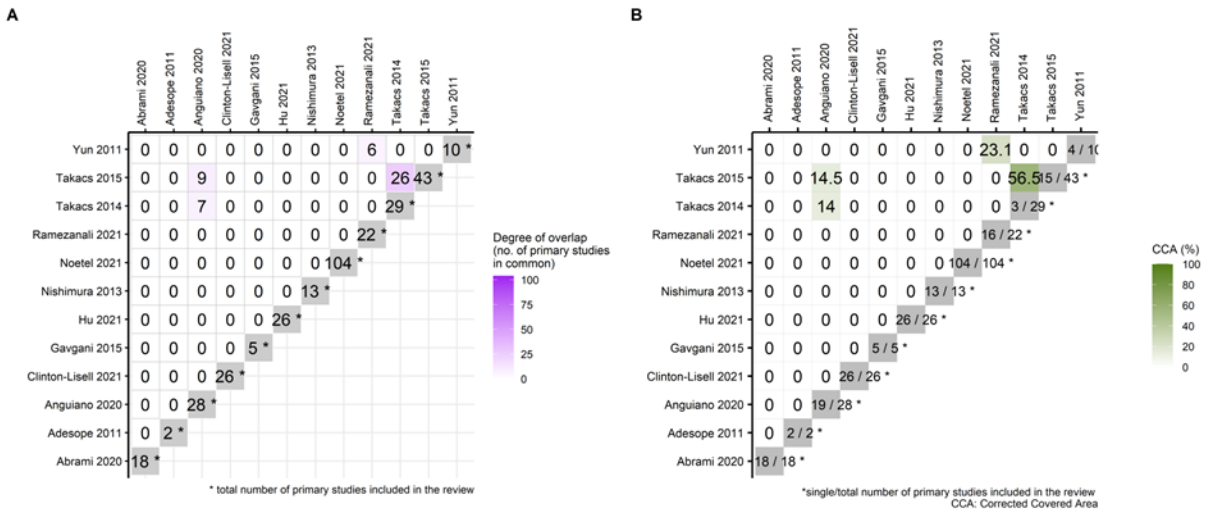
Pairwise Intersection Heatmap

The pairwise intersection heatmap illustrates both the absolute number of overlapping primary studies between each pair of systematic reviews and the corresponding pairwise CCA percentages. These visualizations can be produced using the ccaR() package (Bougioukas et al., 2023) or with an Excel based GROOVE tool from Bracchiglione et al (2022). The map provides not only an overall summary of overlap but also a detailed pairwise comparison.

In Figure 4, Panel A displays the number of shared primary studies between each pair of systematic reviews and Panel B presents the percentage CCA for each pair of reviews. The most notable overlap occurs between Takacs (2015) and Takacs (2014), which share 26 primary studies, corresponding to a high pairwise CCA of 56.5% in Panel B. Additional overlap is observed between Yun (2011) and Ramezanali (2021) (6 shared studies; 23.1%), Takacs (2015) and Anguiano (2020) (9 shared studies; 14.5%), and Takacs (2014) and Anguiano (2020) (7 shared studies; 14.0%). However, most cells in the matrix are zero, indicating no shared primary studies between most review pairs and reinforcing the overall finding of small redundancy.

Figure 4

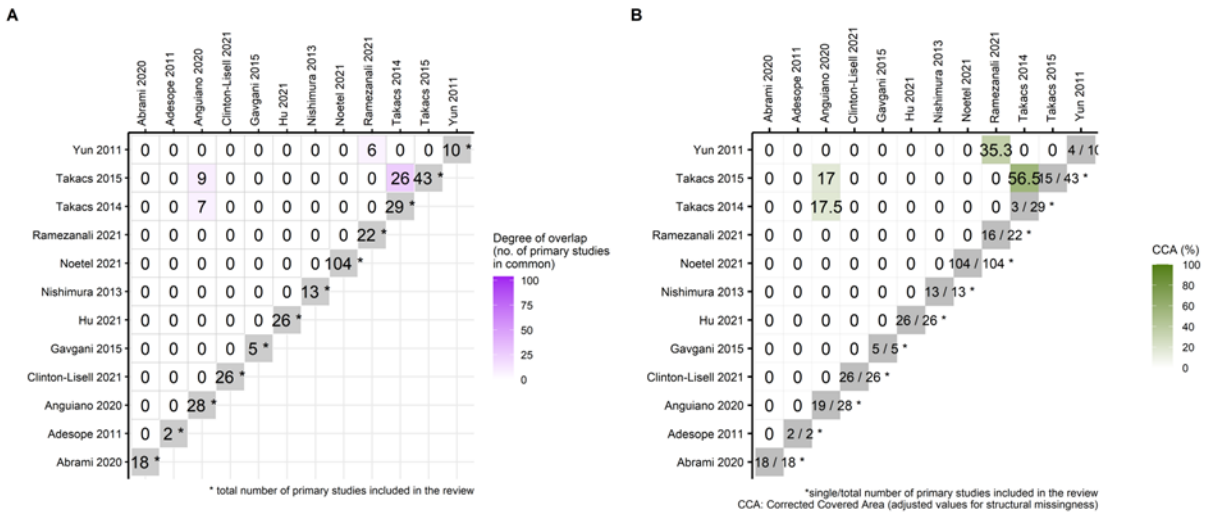
Pairwise Intersection Heatmap



In the adjusted pairwise intersection heatmap (Figure 5), Panel A remains unchanged, as it presents the absolute number of overlapping primary studies between each pair of systematic reviews. Because the percentage CCA values are recalculated to account for structural missingness, Panel B changes after adjustment. The adjusted values are slightly higher and provide a more accurate estimate of overlap.

Figure 5

Adjusted Pairwise Intersection Heatmap



Bougioukas et al. (2021) provide a comprehensive overview of several available visualization methods, including detailed descriptions as well as a discussion of their respective advantages and limitations. Readers interested in a more in-depth examination of overlap visualization techniques are referred to in that paper.

Step 2: Resolving the Overlaps

Once a clear understanding of the extent of overlaps is developed, a strategy for resolving/removing overlaps needs to be deployed to minimize the impact of the validity threat. While there have been recent strides in developing strategies and tools to measure and visualize overlaps, there is far less understanding of how to resolve overlaps. Recent research suggests that the lack of attention to overlaps is problematic and persistent. It is often reported that at least half of published overview studies fail to address overlap, and many other studies only acknowledge it as a threat or quantify it (Gutierrez-Arias et al., 2024; Lunny et al. 2020; Pieper et al., 2014b; Polanin et al., 2017; Pollock et al., 2017;). In this section we describe several prevailing methods for resolving overlaps.

Strategies for Resolution

In a recent review of overviews in healthcare three dominant approaches for resolving overlaps prevailed: computing a new meta-analysis, following a (typically quality-based) algorithm, or author-derived (Gutierrez-Arias et al., 2024). The first of these is extremely inefficient, requiring additional data collection and coding to repeat and combine systematic reviews. Depending on how extensive the revised meta-analysis is, it also could answer a different research question, one that primarily focuses on the summary effect and is therefore not

the same as performing a second-order systematic review. Therefore, we explore the other two options in more depth.

Decision-Based Rules and Algorithms. Various forms of decision algorithms have emerged over time, drawing on early work of assessing statistical independence violations from systematic reviews and meta-analyses. Many early efforts focused on readily available quantitative measures and relied on just one or two criteria for judgement, generally involving a percentage overlap calculation. For example, Sipe and Curlette (1997) used decision rules based on the size of the overlapping studies as well as recency, if sufficient overlap existed, to remove a study (and therefore the overlap). Wilson and Lipsey (2001) proposed a recursive elimination process that removed the largest overlapping studies one-at-a-time until a sufficiently small amount of overlap remained.

However, several authors have also pointed to the value of qualitative information, such as quality indicators, particularly if a systematic review is not exclusively quantitative in focus. The Jadad algorithm (Jadad et al., 1997) was designed to help researchers perform systematic reviews to remove discordant pairs of related research, by considering several qualitative factors including relevance, quality, and selection criteria. Cooper and Koenka (2012) also highlighted other qualitative indicators that could be used to help prioritize removal of studies (e.g., completeness, recency, size, rigor, and source.) Later, Hennesey and Johhson (2020) proposed a decision-tree that incorporates the use of CCA to instigate additional investigation decisions about removal of overlaps.

A few important decisions need to be made when defining a recursive algorithm. The first action is to identify the appropriate comparison source. A researcher could compare the extent that each individual article contains primary studies that overlap with the aggregate field of systematic reviews in the second-order study or individually consider the overlap for all specific pairs of systematic reviews. The next action pertains to the chosen comparison, a specific measure (e.g., percentage or CCA) or measures (e.g., percentage, size, quality, recency) needs to be decided. For comparisons involving just one measure, the decision is determinant. When multiple measures are considered, a framework for valuing each criteria needs to be established. This is often done with a decision-tree (see Jadad et al [1997] or Hennesey & Johnson [2020] for example) or by weighting. Using the chosen measure, an individual study is removed from the analysis, and the process is repeated (e.g., the study with the largest percentage of primary studies overlapping the aggregate field, or one of the two studies with the largest pairwise overlap is removed.) Finally, the last action is to define an appropriate cut-off for the recursive process to end. For example, using percentages, Wilson and Lipsey (2001) suggested a 25% overlap. Newer measures (e.g., using CCA or Adjusted CCA thresholds) could also be used.

Researcher-Based Decisions. Despite all the advances to achieve systematic decision-based algorithms to guide the removal of overlaps, the fact remains that there is a lack of consensus about the proper technique(s) to employ to address overlaps to achieve acceptable levels of statistical independence. Furthermore, there is limited empirical guidance for researchers on how to navigate removal of overlaps efficiently. Consequently, many researchers still present their own methods for resolving overlaps, essentially imposing their own value-judgements into the process (Strech & Tilburt, 2008). While it seems reasonable that researchers would be informed of important considerations within the literature they are reviewing, inconsistency of methods ultimately only limits comparability.

Even though some decision-based algorithms exist, there is still plenty of opportunity for a researcher to insert their own decision into the process, and many do. Some algorithms rely only on an established measure or framework of measures, or defined cut-offs. However, given the relative novelty of conducting second-order reviews, guidelines and standards are not well established or known.

Instructional Application Using Examples from ODDE - Understanding Overlaps

Within ODDE we identified three overview studies that employed a combination of both decision-based algorithms and researcher-based decisions. Each is rooted in the Wilson and Lipsey (2001) method noting a suggested cut-off level of 25% (Sun et al., 2025; Martin et al., 2022; Tamim et al., 2011). Martin et al. (2022) describe a sensitivity analysis to compare the final analysis with less restrictive cut-off levels (e.g., 50%). Sun et al. (2025) used a 50% cut-off given a limited sample size.

As illustrated in Figure 3, some studies exceed the 50% cut-off level (i.e., there are more primary studies appearing in multiple papers than in just one). In this case Takacs et al. (2014) has the highest percentage of primary studies overlapping with other reviews, so it is removed. As Tamim et al. (2001) points out, the recursive process may take several iterations. Further exploration of Figure 1 shows Takacs et al. (2014) and Takacs et al. (2015) are highly aligned. The removal of just one of these studies limits the percentage of overlaps in the remaining reviews, and no other removal was required. As a result, all estimations of effects would need to be recalculated using the smaller dataset.

The process described above utilized percentage of overlap, but one could easily conceptualize other rules including the use of pairwise CCAs and Adjusted CCAs; however, Pieper et al (2014) and others (e.g., Lunny [2020]) cautioned against using CCA as an absolute cutoff and instead have advocated for a more qualitative-based decision. Ultimately, further research into the use of CCA-type measures is needed to assess the sensitivity of the measure to different overlap scenarios.

In the Sun et al. (2025) example, the results are the same. Pairwise comparisons would highlight the significant overlap between Takacs (2014) and Takacs (2015), and measures of recency and size would dictate removal of the 2014 piece. Alternatively compared to the field, the most overlap occurs with the 2014 article, and it is therefore first in line for removal. It's easy to hypothesize a scenario where one article overlaps significantly with the field but doesn't have a large pairwise overlap with any other article.

Discussion

Second-order meta-analyses are an important tool for combining first-order meta-analyses to examine various related interventions, contextual factors, and study populations. They are useful for synthesizing and updating evidence as innovations change how interventions are conceptualized and used. Yet, they also present new challenges, like managing overlaps, for researchers to navigate. Standardizing practice is a necessary step for ensuring findings from these studies are useful. In our experience, addressing overlaps in second-order studies is a daunting challenge, with much of the research residing in unfamiliar fields.

For practitioners working in ODDE, several implications for practice and future research can be drawn from this synthesis. The ODDE field evolves quickly with new technologies generating widespread use within a short period of time. The recent example of generative artificial intelligence is a good example that continues to transform education in all contexts: a small finite time frame, but many meta-analyses. Thus, within ODDE many second-order studies are likely to encounter considerable overlap given narrow time frames. Encouraging consistent practices is important for promoting the great value of these second-order studies. Within ODDE, significant challenges of transparency are already apparent (Zawacki-Richter et al, 2025). In Sun et al. (2025), a study of only 15 meta-analyses, 20% of the included meta-analyses did not have adequate data to include in the overlaps assessment.

To adequately address overlaps, researchers need to first thoroughly understand the overlaps through calculation and visualization. The researcher's first goal should be to generate a citation matrix and produce a pairwise heat map and stacked bar chart. With these outputs the researcher can adequately describe the range, frequency, and distribution of pairwise CCA comparisons as well as the overall CCA. However, some important ODDE questions consider impacts across generations (Moore & Kearsley, 2011). For ODDE studies that consider questions spanning different generations (e.g., Martin et al., 2022), use of adjusted CCA in the overlap quantification is an important consideration. However, extending calculations beyond CCA to consider more refined calculations including wCCA appears to be difficult within ODDE, given inconsistencies in reporting of meta-studies. Promoting greater consistency of reporting practices for meta-analyses with an aim to also utilize these meta-analyses for second-order studies will enhance transparency and enable more refined research.

The PRIOR guidelines (Gates, 2022) have sought to align researchers in how to characterize and report on overlaps (Pollock et al., 2019). However, little prescription is offered to manage overlaps beyond "Describe the extent of primary study overlap across the included systematic reviews." (Gates, 2022, data supplement). Given recent advances to efficiently visualize and quantify overlaps, the empirical derivation of best practices presents an emerging research need (Bracchiglione et al., 2026), with empirical comparisons of resolution strategies remain a particularly important avenue for future research. Overall, ODDE researchers should clearly assess and report overlaps and transparently justify the decision to retain the overlaps without adjustment or removal.

The present work represents a methodological commentary, helping to import tools and strategies from other fields in an effort to advance the ODDE field into better practice and encourage more empirical research in this area. Broadly, there has been limited empirical research on quantifying, visualizing, and resolving overlaps (Gutierrez-Arias et al., 2024; Lunny, 2020). Even in healthcare, where much of the research resides, there is a dearth of empirical evidence depicting the impact of strategies used to limit violations of statistical independence in second-order studies. Such a dearth of guidance limits not only the guidance to move forward, but also negatively impacts the validity inferences based upon existing overview studies, as it is unclear the extent that failing to address overlap can have on robustness of reported findings. As Schmidt and Oh (2013) suggest, more extensive simulation work (akin to Tracz, Elmore, and Pohlmann [1992]) is needed to inform decision-making around overlaps in second-order studies. In our judgement, the ODDE context presents an important context for providing empirical research to help justify and dictate best practices. . As the number of ODDE second-order reviews grows, systematic reviews should be conducted to describe and examine how these

strategies are applied in wider practice. Additionally, future research should use simulation studies to empirically evaluate the performance of different overlap management strategies across varying overlap scenarios.

Limitations

The illustrative nature of this study introduces several limitations. Our application draws on a single ODDE overview in which the overall level of overlap is relatively low, limiting the opportunity to examine different approaches under severe overlap conditions. The application of wCCA also revealed a practical constraint, as inconsistent reporting of primary-study sample sizes made the measure difficult to implement fully. This paper is intended to illustrate and discuss available approaches rather than to systematically and empirically compare their performance, which is a necessary next step for formulating best-practices. Thus, the running example is intended to be demonstrative and should not be interpreted as evidence that one resolution strategy or decision rule is better than another.

References

Anguiano, C. J. (2020). *Effects of multimedia-enhanced storybooks on young childrens' vocabulary and comprehension knowledge: A meta-analysis*. (Publication No. 27955486) [Doctoral dissertation, Washington State University]. *ProQuest Dissertations & Theses Global*.

Becker, L. A., & Oxman, A. D. (2008). Overviews of reviews. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions: Cochrane book series.* John Wiley。 https://doi.org/10.1002/9780470712184.ch22

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* John Wiley & Sons. https://doi.org/10.1002/9780470743386

Borokhovski, E., Tamim, R., & Pickup, D. (2022). *Technology application in teaching and learning: Second-order review of meta-analyses.* Commonwealth of Learning. https://doi.org/10.56059/11599/4069

Bougioukas, K. I., Diakonidis, T., Mavromanoli, A. C., Haidich, A.-B. (2023). ccaR: a package for assessing primary study overlap across systematic reviews in overviews. *Research Synthesis Methods, 14*, 443-454. https://doi.org/10.1002/jrsm.1610

Bougioukas, K. I., Vounzoulaki, E., Mantsiou, C. D., Savvides, E. D., Karakosta, C., Diakonidis, T., Tsapas, A., & Haidich, A.-B. (2021). Methods for depicting overlap in overviews of systematic reviews: An introduction to static tabular and graphical displays. *Journal of Clinical Epidemiology*, *132*, 34–45. https://doi.org/10.1016/j.jclinepi.2020.12.004

Bracchiglione J., Meza, N., Bangdiwala, S. I., Niño de Guzmán, E., Urrútia, G., Bonfill, X., & Madrid, E. (2022). Graphical Representation of Overlap for OVErviews: GROOVE tool. *Research Synthesis Methods,13*(3): 381–388. https://doi.org/10.1002/jrsm.1557

Bracchiglione, J., Meza, N., Pérez-Carrasco, I., Vergara-Merino, L., Madrid, E., Urrútia, G., & Bonfill Cosp, X. (2023). A methodological review finds mismatch between overall and pairwise overlap analysis in a sample of overviews. *Journal of Clinical Epidemiology*, *159*, 31–39. https://doi.org/10.1016/j.jclinepi.2023.05.006

Bracchiglione J., Meza, N., Pieper, D., Lunny, C., Vargas-Peirano, M. Vicuña, J., Briceño, F., Garnham Parra, R., Pérez Carrasco, I., Urrútia, G., Bonfill, X., & Madrid, E. (2026). Impact of matrix-construction assumptions on quantitative overlap assessment in overviews: A meta-research study. *Research Synthesis Methods, 17*, 348-364. https://doi.org/10.1017/rsm.2025.10056

Cooper, H., & Koenka, A. C. (2012). The overview of reviews: Unique challenges and opportunities when research syntheses are the principal elements of new integrative scholarship. *American Psychologist, 67*(6), 446–462. https://doi.org/10.1037/a0027119

Gates, M., Gates, A., Pieper, D., Fernandes, R. M., Tricco, A. C., Moher, D., Brennan, S. E., Li, T., Pollock, M., Lunny, C., Sepúlveda, D., McKenzie, J. E., Scott, S. D., Robinson, K. A., Matthias, K., Bougioukas, K. I., Fusar-Poli, P., Whiting, P., Moss, S. J., & Hartling, L. (2022). Reporting guideline for overviews of reviews of healthcare interventions:

Development of the PRIOR statement. *BMJ, 378*, e070849. https://doi.org/10.1136/bmj-2022-070849

Gnanasanjevi, G., Balasubramaniam, P., Sriram, N., Lakshmi, S., Gangai Selvi, R., & Suresh, J. (2025). Impact of open and distance learning on agricultural education: A systematic review and bibliometric analysis. *Plant Science Today, 12*(2). https://doi.org/10.14719/pst.7505

Gutierrez-Arias, R., Pieper, D., Lunny, C., Torres-Castro, R., Aguilera-Eguıa, R., Oliveros, M.-J., & Seron, P. (2024). Only half of the authors of overviews of exercise-related interventions use some strategy to manage overlapping primary studies–a metaresearch study. *Journal of Clinical Epidemiology, 170*, 111328. https://doi.org/10.1016/j.jclinepi.2024.111328

Hennesey, E. A., & Johnson, B. T. (2020). Examining overlap of included studies in meta-reviews: guidance for using the corrected covered area index. *Research Synthesis Methods, 11*,134-145. https://doi.org/10.1002/jrsm.1390

Jadad, A. R., Cook, D. J., & Browman, G. P. (1997). A guide to interpreting discordant systematic reviews. *Canadian Medical Association Journal, 156*(10), 1411-1416.

Lunny, C., Brennan, S. E., McDonald S., & McKenzie, J. E. (2017). Toward a comprehensive evidence map of overview of systematic review methods: paper 1—purpose, eligibility, search and data extraction. *Systematic Reviews*, 231. https://doi.org/10.1186/s13643-017-0617-1

Lunny, C., Brennan, S. E., McDonald S., & McKenzie, J. E. (2018). Toward a comprehensive evidence map of overview of systematic review methods: paper 2—risk of bias assessment; synthesis, presentation and summary of the findings; and assessment of the certainty of the evidence. *Systematic Reviews*, 159. https://doi.org/10.1186/s13643-018-0784-8

Lunny, C., Brennan, S. E., Reid, J., McDonald S., & McKenzie, J. E. (2020). Overviews of reviews incompletely report methods for handling overlapping, discordant, and problematic data. *Journal of Clinical Epidemiology, 118*, 69-85. https://doi.org/10.1016/j.jclinepi.2019.09.025.

Lunny, C., Pieper, D., Thabet, P., & Kanji, S. (2021). Managing overlap of primary study results across systematic reviews: practical considerations for authors of overviews of reviews. *BMC Medical Research Methodology, 21*(1), 140. https://doi.org/10.1186/s12874-021-01269-y

Martin, F. (2025). *Systematic reviews in educational research: A practical guide for research synthesis.* Thousand Oaks, CA: Sage Publications, Inc.

Martin, F., Sun, T., Westine, C. D., Ritzhaupt, A. D. (2022). Examining research on the impact of distance and online learning: A second-order meta-analysis study. *Educational Research Review, 26*, 100438.

Moore, M. & Kearsley, G. (2011). *Distance education: A systems view* (3rd edition). Cengage Learning.

Pieper, D., Antoine, S. L., Mathes, T., Neugebauer, E. A., & Eikermann, M. (2014). Systematic review finds overlapping reviews were not mentioned in every other overview. *Journal of Clinical Epidemiology, 67*(4), 368-375. https://doi.org/10.1016/j.jclinepi.2013.11.007

Pollock, A., Campbell, P., Brunton, G., Hunt, H., & Estcourt, L. (2017). Selecting and implementing overview methods: implications from five exemplar overviews. *Systematic reviews, 6*(1), 145. https://doi.org/10.1186/s13643-017-0534-3

Pollock, M., Fernandes, R. M., Pieper, D., Tricco, A. C., Gates, M., Gates, A., & Hartling, L. (2019). Preferred Reporting Items for Overviews of Reviews (PRIOR): A protocol for development of a reporting guideline for overviews of reviews of healthcare interventions. *Systematic reviews, 8*(1), 335. https://doi.org/10.1186/s13643-019-1252-9

Polanin, J. R., Maynard, B. R., & Dell, N. A. (2017). Overviews in Education Research: A Systematic Review and Analysis: A Systematic Review and Analysis. *Review of Educational Research*, *87*(1), 172-203. https://doi.org/10.3102/0034654316631117

Ramezanali, N., Uchihara, T., & Faez, F. (2021). Efficacy of multimodal glossing on second language vocabulary learning: A meta-analysis. *TESOL Quarterly*, *55*(1), 105-133. https://doi.org/10.1002/tesq.579

Schmidt, F. L., & Oh, I. S. (2013). Methods for second order meta-analysis and illustrative applications. *Organizational Behavior and Human Decision Processes, 121*(2), 204–218. https://doi.org/10.1016/j.obhdp.2013.03.002

Sipe, T. A., & Curlette, W. L. (1997). A meta-synthesis of factors related to educational achievement: A methodological approach to summarizing and synthesizing metaanalysis. *International Journal of Educational Research, 25*, 583–698. https://doi.org/10.1016/S0883-0355(96)80001-2

Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What One Hundred Years of Research Says About the Effects of Ability Grouping and Acceleration on K-12 Students' Academic Achievement: Findings of Two Second-Order Meta-Analyses. *Review of Educational Research*, *86*(4), 849–899. http://www.jstor.org/stable/44668238

Strech, D. & Tilburt, J. (2008). Value judgements in the analysis and synthesis of evidence. *Journal of Clinical Epidemiology, 61*(6), 521-524. https://doi.org/10.1016/j.jclinepi.2008.01.001

Sun, T., Wu, T., Martin, F., & Westine, C. D. (2025). Effects of multimedia-based education on learning outcomes: An evidence gap map analysis and second order meta-analysis. *Journal of Educational Multimedia and Hypermedia, 32*(1), 57-95. https://doi.org/10.70725/798356qwfmej

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2014). Can the computer replace the adult for storybook reading? A meta-analysis on the effects of multimedia stories as compared to sharing print stories with an adult. *Frontiers in psychology*, *5*, 1366.

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks: A metaanalysis. *Review of Educational Research*, *85*(4), 698-739.

Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research, 81*(1), 4-28. https://doi.org/10.3102/0034654310393361

Tracz, S. M., Elmore, P. B., & Pohlmann, J. T. (1992). Correlational meta-analysis: Independent and nonindependent cases. *Educational and Psychological Measurement, 52*, 879-888. https://doi.org/10.1177/0013164492052004007

Wilson, D. B., & Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods, 6*, 413–429. https://doi.org/10.1037/1082-989x.6.4.413

Yang, H. L. & Xiang, J. J. (2024). Mobile learning and students' academic achievement: A second-order meta-analysis study. *Journal of Educational Computing Research, 62*(4), 741-763. https://doi.org/10.1177/07356331241226589

Ying, X., Bougioukas, K. I., Pieper, D., & Mayo-Wilson, E. (2025). Weighted corrected covered area (wCCA): A measure of informational overlap among reviews. *Research Synthesis Methods*, 702-708. https://doi.org/10.1017/rsm.2025.19

Yun, J. (2011) The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning, 24*(1), 39-58. https://doi.org/10.1080/09588221.2010.523285

Zawacki-Richter, O., Cefa, B., & Bai, J. Y. (2025). *Towards reproducible systematic reviews in Open, Distance, and Digital Education—An umbrella mapping review.* Review of Education, 13(1), e70031. https://doi.org/10.1002/rev3.70031