journal of mixed methods studies

Issue 12, 2025
Journal of Mixed Methods Studies / JOMES

Research Article

A Methodological Protocol for Analyzing Dyadic Phenomena: The Cross-Network Informational Analysis (CNIA)

Donizete Ferreira Beck Marcos Ferasso

To cite this article:
Beck, D., & Ferasso, M. (2025). A methodological protocol for analyzing dyadic phenomena: The cross-network informational analysis (CNIA). *Journal of Mixed Methods Studies, 12*, 1–21. https://doi.org/10.59455/jomes.51

Article Info: *Received*: 12 December 2024 *Revised*: 4 February 2025 *Accepted*: 8 March 2025

Abstract
This paper proposes a methodological protocol for analyzing dyadic phenomena based on two (or more) categorical variables by integrating qualitative and quantitative methods. A novel methodological protocol based on the MMR Paradigm was proposed to meet the need for adequately exploring dyadic phenomena by a methodological integration that increases the robustness and in-depth exploration of findings. This protocol consisted of Research Synthesis, Classification, Information Network Analysis, and Exploratory Data Analysis, finalized by a meta-inference. Its application suits empirical research, literature reviews, and other theoretical studies. The main contribution of the CNIA protocol to MMR is its application for theoretical research dealing with dyadic phenomena under the MMR Paradigm. A practical application and methodological suggestions were presented.

Keywords: *dyadic phenomena; categorical variables; mixed methods research paradigm; informational network analysis; research synthesis*

A Methodological Protocol for Analyzing Dyadic Phenomena: The Cross-Network Informational Analysis (CNIA)

Introduction

In any field of study, there are many ways of conducting theoretical studies, such as Bibliometrics, Research Synthesis (RS), Systematic Literature Review (SLR), and meta-analysis, only to name the most common. Some of these methods are more used than others, remaining preferred theoretical and empirical studies, varying from one field of study to another. Nevertheless, researchers have the practicality of choosing one given method already assumed by scholars and following its steps to achieve results. Less commonly, researchers use methods in a combined way, such as bibliometrics and Network Analysis (NA) or SLR and bibliometrics.

Dyadic phenomena (DPs) have been more commonly studied in some fields like Administrative Sciences, Social Sciences, Politics, Psychology, and Medicine (Korsgaard et al., 2015; Goldsmith et al.2017; Lyons & Lee, 2018; Kim et al., 2020). However, the available methodologies have not fully responded to the needs of such an approach (Krasikova & LeBreton, 2012). For instance, although various categorical variables (CVs) can be in-depth studied through heterogeneous information NA (Sun et al., 2012), a qualitative analysis could improve interpretations by getting integrated with the quantitative results.

Few methodologies propose an *integrated approach* for gaining a deeper and broader understanding of the phenomena under study (Åkerblad, Seppänen-Järvelä & Haapakoski, 2021; Nooraie et al., 2020). This integration is pivotal in *Mixed Methods Research* (MMR) when analyzing empirical findings from quantitative and qualitative strands and integrating the different inferences into a meta-inference (Teddlie & Tashakkori, 2009). Among the benefits of MMR is that this methodological approach strengthens the inferences quality (Tashakkori, Johnson & Teddlie, 2020). This manuscript departs from the methodological gap in addressing the need for an in-depth exploration of DPs and the MMR precepts by proposing a new integrative methodological protocol. Thus, this methodological paper aims to propose a methodological protocol for analyzing DPs based on two (or more) different CVs by integrating qualitative and quantitative methods.

In this protocol, named *Cross-Network Information Analysis* (CNIA), we departed from the robustness proportioned by the MMR and its integration to strengthening the power of explanation for CVs in DPs (Krasikova & LeBreton, 2012; Tashakkori et al., 2020). The CNIA protocol incorporates two strands: one qualitative - comprising RS, classification, and inferences; and another quantitative - comprising Informational NA, Exploratory Data Analysis (EDA), and inferences. The integration of inferences of each strand is performed at the final analysis (i.e., meta-inference).

The originality of this methodological protocol contributes to scholars in many ways. First, to the best authors' knowledge, no methodological proposal is available for analyzing DPs with two or more CVs. Second, a novel integrative protocol for analyzing CVs of the DP was proposed and applied. Third, this methodology has qualitative and quantitative strands that increase the *rigor* and *reliability* of results. Fourth, this research offered a new qualitative-quantitative procedure for analyzing data via Informational NA precepts, including a classification that preceded the quantitative analysis. Fifth, more significant results are achieved when integrating the inferences into the meta-inference

Methodological and Conceptual Background

In this section, the methodological concepts of science and data analysis, RS, NA, and EDA are revisited. These four methodological backgrounds help understand the procedures used for building CNIA.

Science and Data Analysis

Advancing the frontiers of knowledge and understanding of phenomena are the driver forces of science literature. In this way, scholars make empirical or theoretical advancements. In all cases solid and reliable *information* is used for building scientific knowledge, which is done by performing rigorous methods and analyses. For this reason, *data analysis* plays a critical role as an optimal tool for the accuracy and objectivity of science research (Hair et al., 2019). The deepened understanding of data analysis allows scientists to work with theoretical constructs through propositions and hypotheses at the abstract and empirical levels, which implies in analyzing distinct data and variables (Laudan, 1981; Hair et al., 2019). The main types of *variables for data analysis* are (Agresti, 2002; Hair et al., 2019): CVs; quantitative variables; response variables; and explanatory variables.

CVs, also known as qualitative variables and nonmetric variables, are defined as variables "with values that serve merely as a label or means of identification" (Hair et al., 2019, p. 473). CVs are also elements of information networks, which are part of DPs (Krasikova & LeBreton, 2012; Sun et al., 2012). CVs can be nominal or ordinal, and dichotomous or multichotomous (Agresti, 2002; Hair et al., 2019). The CNIA protocol intends to analyze information networks of DPs. Information networks comprise two or more CVs. Although NA metrics are quantitative, the CVs are the informational input of networks. Moreover, networks are critical for analyzing DPs. Thus, this section underlined the relevance of rigor in analyzing quantitative and qualitative data for scientific development. Furthermore, this section presents the concepts and main types of variables for data analysis.

Research Synthesis

RS is a replicable systematic technique for exploring and analyzing findings from previous research (Mosteller & Colditz, 1996; Atkinson et al., 2015). Initially used in Life Sciences, other sciences fields have used RS, such as Linguistics (Norris & Ortega, 2000), Teaching and Education (Minner, Levy & Century, 2010), Managerial Sciences (Beck & Ferasso, 2022), and Engineering (Cordray, Harris & Klein, 2009). Also, RS has gained space mainly among Social Sciences scientists after the post-World War II (Chalmers, Hedges & Cooper, 2002; Hedges & Cooper, 2009).

Furthermore, RS is also recommended to complement literature review studies due to an integrative perspective of empirical research for creating generalizations (Cooper et al., 2019). Furthermore, RS allows a *deeper understanding of the findings* of retrieved research papers thanks to the qualitative approach that analyzes and integrates the findings from selected studies, and the interpretation resulting from the found evidence (Cooper et al., 2019; Cooper, 2015; Suri, 2011).

Thus, RS was included in CNIA by considering the RS of MMR (Voils et al., 2008; Sandelowski et al., 2012). In this perspective, qualitative and quantitative studies are previously examined throughout RS by aggregation. RS by aggregation comprises assimilating the findings to have the same structure or connection between two aspects or features of the phenomenon under study (Voils et al., 2008; Sadelowski et al., 2012). The ultimate goal of any RS is to produce a *cumulative integration of results* from previous research aiming at the quality of the information resulting from these analysis processes (Heyvaert, Maes & Onghena, 2013). Then, the cumulative integration of results is used for the next stage, the codification.

Codification is one of the critical aspects of RS. A coding scheme, generated by the researcher, depends on his knowledge on the subject under study for deciding to which category an item pertains. The categories are generated by code forms or code book, making the coding procedure more accurate. The main purpose of the coding scheme is to provide further *classification of items into categories* (Stock, 1994), which is useful for dealing with CVs. The researcher must define procedures for evaluating coding decisions and avoiding errors, such as ambiguities in the judgment process, coder bias, and coder mistakes (Orwin, 1994). For non-numerical information, classification occurs according to the *variations of characteristics* found across the analyzed studies, which is part of the coding convention. The reliability of coders and coding procedures are checked through revisions of coded results by other coders following a coding scheme protocol (Stock, 1994). Therefore, RS provides the *principles to synthesizing data and classifying data items into CVs* for CNIA.

Network Analysis

In CNIA, *NA* is helpful for analyzing the pattern of complex interplay among - networked elements - from one or more *CVs*. *Network* is "a collection of points joined together in pairs by lines" (Newman, 2018, p. 1). In scientific terms, points within networks are known as *nodes* and the lines are *edges*. In this way, *NA* "is useful to understand the pattern of interactions of nodes and edges within a system" (Beck & Ferasso, in press, p. 6). The systemic relationships analyzed in networks can be made from many science fields (Borgatti et al., 2009; Newman, 2018), and the main network *types* are: *technological*, e.g., the physical internet system that links computers by cables; *social*, e.g., social media, teams, groups of people, and firms; *biological*, e.g., neural networks, and food web of predator-prey relationships; and *informational*, e.g., texts, speeches, world wide web, literature, and bibliometric data.

Furthermore, it is crucial to understand the *type of ties* since they represent the *DPs* represented in the network (Borgatti et al., 2009). DPs occur when at least two elements are interrelated due to some nature or determined reason. According to Borgatti et al. (2009), the main typology of ties in network analyses (Figure 1) are variables based on similarity (e.g., attribute, location, and membership similarities), social relations (e.g., role-based, kinship-based, affective-based, and cognitive-based relationships), interactions (attitudes and actions, such as having sex with, advincing, helping, and harming someone), and flows (e.g., flows of information, resources, and beliefs).

Figure 1

A Typology of Ties Studied in Network Analysis

Similarities			Social Relations				Interactions	Flows
Location	**Membership**	**Attribute**	**Kinship**	**Other role**	**Affective**	**Cognitive**	e.g.,	e.g.,
e.g.,	e.g.,	e.g.,	e.g.,	e.g.,	e.g.,	e.g.,	Sex with	Information
Same spatial and temporal space	Same clubs	Same gender	Mother of	Friend of	Likes	Knows	Talked to	Beliefs
	Same events	Same attitude	Sibling of	Boss of	Hates	Knows about	Advice to	Personnel
				Student of	etc.	Sees as happy	Helped	Resources
	etc.	etc.		Competitor of			Harmed	etc.
						etc.	etc.	

Note. From "Network Analysis in the Social Sciences," by S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, 2009, *Science*, *323*(5916), p. 894. Copyright 2009 by American Association for the Advancement of Science.

In this way, networks can be directed or undirected. *Directed networks* have edges pointing out a DP originated from one node to another, and *undirected networks* represent reciprocity regarding the DP. For instance, the food web network is a directed network because one species eats another species (species are the nodes, and eating represents the edges); and friendship is an undirected network because friendship is reciprocal among people (people are nodes, and friendships are the edges). In essence, *edges* represent directed or undirected DPs that connect a *node* to another, and *nodes* can be concrete or abstract things from a variety of network types and fields.

The fundamental metrics for NA can be applied for all network types. However, each metric will reveal specific aspects of the networks, and for this reason, the chosen variable should be aligned with the respective purpose of the analysis (Borgatti et al., 2009; Newman, 2018). The most used metrics in NA are centrality measures, since the position of a node within a network "determines in part the opportunities and constraints that it encounters, and in this way plays an important role in a node's outcomes" (Borgatti et al., 2009, p. 894). Also, other important metrics in NA are reciprocity, transitivity, homophily, and similarity.

Centrality is based on the idea that some nodes play a critical role in the network (Borgatti et al., 2009; Newman, 2018), and this measure quantifies the importance of a node in a given network. In general, centrality reveals the *outstanding* nodes in networks. The three main measures of Centrality have been 'Degree Centrality' (DC), 'Closeness Centrality' (CC), and 'Betweenness Centrality' (BB) (Bastian, Heymann & Jacomy, 2009; Newman, 2018). Other important centrality measures are 'PageRank', 'Eigenvector Centrality', and 'Katz Centrality' (Newman, 2018).

DC is basically "the number of edges connected to it" (Newman, 2018, p. 159), which reveals the *nodes* with more power of influence or prestige in networks. In directed networks, DC can be *in-degree* (i.e., quantity of edges inward a node) and *out-degree* (i.e., quantity of edges outward a node). In undirected networks, centrality reveals the quantity of edges connected to a node. *CC* "measures the mean distance from a node to other nodes" (Newman, 2018, p. 170), and thus, it identifies the nodes with the - shortest mean distance - to other nodes in the network. Nodes with higher CC have better accessibility and fast influenceability to other nodes. DC and CC have generally been positively correlated (Newman, 2018). *BC* "measures the extent to which a node lies on paths between other nodes" (Newman, 2018, p. 173), and thus, it identifies which nodes are in the pathway between other nodes. High BC can indicate that a node has informational control within a network and if this node is removed from the network, the edges and distance between the nodes will hugely change (Newman, 2018).

In *Information NA*, homogeneous and heterogeneous pieces of information can be considered (Yu, Han & Faloutsos, 2010; Sun et al., 2012). Information NA has been used as tool for: social-media analysis (Schmitt et al., 2018); multiple subfields of social, economic, environmental, and political research (Jacobs & Cramer, 2017; Barbi & Prataviera, 2019); and medical, biological, and pharmaceutical research (Wang et al., 2016). On the one hand, *Homogeneous Information Networks* have only one type of information in the network, e.g., the nodes in a given network are only articles. On the other hand, *Heterogeneous Information Networks* have two or more types (or clusters) of information (Sun et al., 2012; Shi et al., 2016; Shi et al., 2017; Xie et al., 2021), e.g., the nodes could be theoretical concepts, phenomena, and science literature publications connected in a given network. In this sense, NA has also been useful for representing qualitative data (Canché, 2022).

Two examples of Heterogeneous Information Network are: (1) In the study of Beck and Ferasso (in press), in which two clusters of nodes were analyzed: 'Sustainable Development Goals' and 'Stakeholder Capitalism Literature'; and (2) in Wang et al. (2019) that investigated 'Genes' and 'Ovarian Cancer'. Although the two examples aforementioned have only two *CVs* in each study, research can utilize more than two variables if demanded by the research purpose (Borgatti et al., 2009; Sun et al., 2012; Newman, 2018). Shi et al. (2016) argues that *Heterogeneous Information Network Analysis* allows creating a new way of developing data mining, and fusing more complex information with multiple semantics. Thus, it allows "merging information from heterogeneous sources with differing conceptual, contextual and typographical representations" (Shi et al., 2017, p. 11).

Therefore, analyzing *Heterogeneous Information Networks* is a useful tool for analyzing informational complex structures with diverse semantics, since it "contains abundant knowledge about relationships among objects" (Sun et al., 2012, p. 2023). Thus, the complexity of heterogeneous information networks can be addressed by fused analysis under MMR Paradigm (Nooraie et al., 2020). For this reason, the *Heterogeneous Information Network* is the optimal network type to be performed in CNIA.

Exploratory Data Analysis

EDA provides excellent means for identifying the properties and analyzing the distribution of a sample data. Thus, EDA analyzes the behavior of the data elements for a determined variable. The main sample data properties revealed in EDA are the means, medians (i.e., second quartile), modes, first and third quartiles, and outliers (Tukey, 1977; Hair et al., 2019). Identifying outliers is crucial in EDA. *Outliers* "are elements having significantly different behavior under a variable in a dataset" (Beck & Ferasso, in press, p. 8). In networks, a node outlier may be "shows irregularity in its structure within its locality", an edge outlier may connect "disparate communities of nodes", it can be an inference based on EDA analysis of a Network metric, or any other abnormal behavior of network data (Aggarwal, 2017, p. 25).

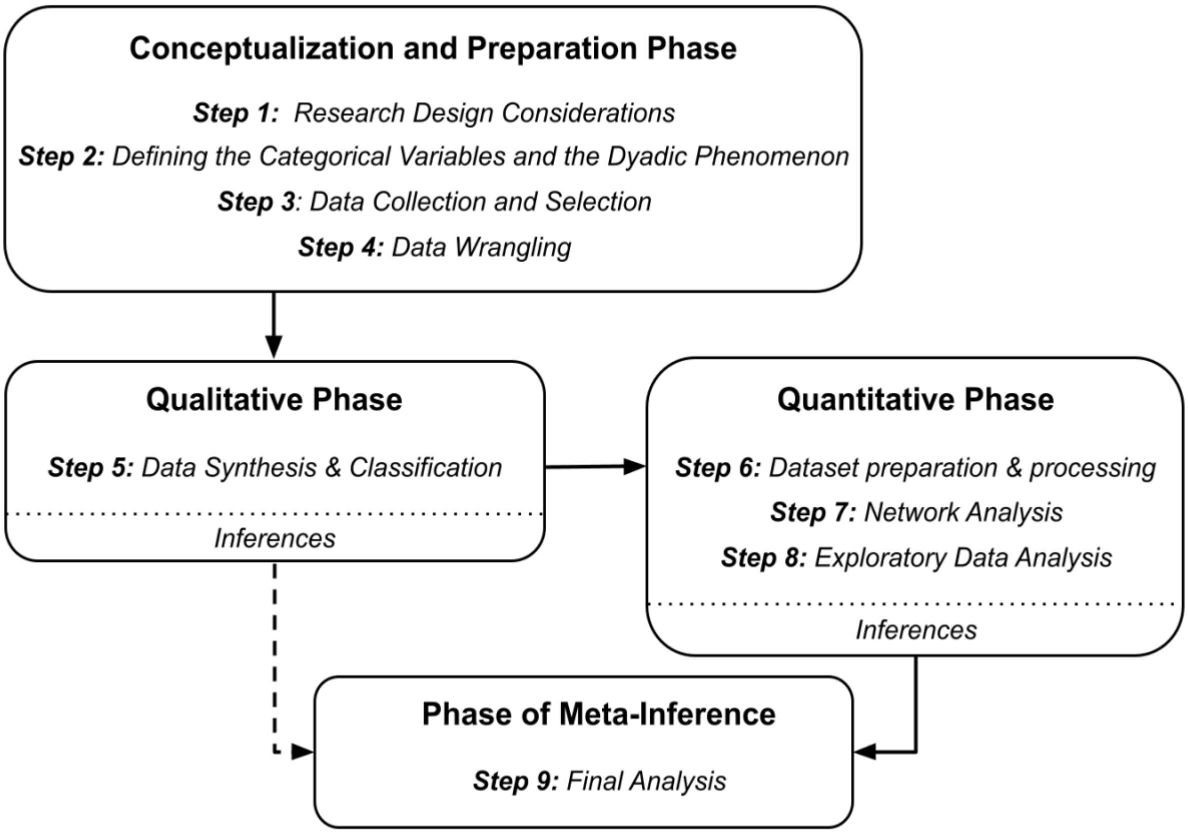
Furthermore, boxplot is an efficient EDA tool for visualizing data in overview and plotting outliers (Nuzzo, 2016) by considering the interquartile range (IQR) and whisker limits of Tukey (1977) or Altman (1991). The Altman's whiskers are only recommended for samples bigger than forty elements because they extend the whiskers limits to the 5th and 95th percentiles, and thus, considering only 90% of the data and more elements as outliers (Spitzer et al., 2014; Nuzzo, 2016) In other words, EDA has also been applied on sample data in NA for identifying outliers, properties, and network behavior (Aggarwal, 2017; Beck & Storopoli, 2021; Beck & Ferasso, in press). Therefore, the metrics of NA in CNIA can be carefully scrutinized by EDA as well as identifying the outliers in the networks.

Guidelines for Cross-Network Informational Analysis

The CNIA protocol was conceived into four phases, starting with a conceptualization and preparation phase, as depicted in Figure 2. Then, qualitative and quantitative phases are performed, and the final phase is the meta-inference.

Figure 2

Cross-Network Informational Analysis' General Guidelines: The Phases and Steps



Each of the four phases are discussed in detail in the next sections.

Conceptualization and Preparation Phase

Step 1: Research Design Considerations

Initially, researchers should decide if CNIA is appropriate to their intended research. There are three primary considerations here. First, the reasons for conducting the research should be clear and concise on (Tashakkori et al., 2020): (a) the social, practical and theoretical relevance; and (b) researchability of the data, theory, and sources; and (c) a justifiable gap based on a DP. Moreover, in CNIA, the analyzed purpose should aim to explore the DPs of two or more data elements or variables.

Second, as CNIA is based in part on NA, the research purpose should be related to analyzing the characteristics and behavior of the variables and data among them. In other words, the properties of the variables and data in a given network are relevant to explain the studied phenomenon stated in the research purpose. For instance, in the Physical Sciences, scholars have analyzed the "universal characteristics of nonrandom networks", while in Social Sciences, they have focused on the "variation in structure across different groups or contexts, using these variations to explain differences in outcomes" (Borgatti et al., 2009, p. 894). Therefore, it could be viable to apply CNIA if the research purpose aims to explore the *universality* or *variability* of the variables and data.

Third, data size matters. On the one hand, the data size should not be ample to the extent that the researcher can not synthesize and classify the data. On the other hand, the data size should not be too small to the extent that the scientific rigor is denied. For instance, it cannot correctly execute EDA in variables with less than five elements (Nuzzo, 2016). Therefore, the data size should be considered when deciding if CNIA or other method protocol will be performed.

Step 2: Defining the Categorical Variables and the Dyadic Phenomenon

The CVs should be chosen by considering the *research question* and the *DP*. The research purpose indicates the type of data to be used in the CNIA, which should be CVs. For instance, the research purpose of the study of Beck and Ferasso (in press, p. 1) was to explore "how Stakeholder Capitalism can contribute to global governance to achieve all of the 17 SDGs". They chose the 'scientific literature' on stakeholder capitalism to represent one CV, and the 'Sustainable Development Goals' for the other CV. In that example, the DP is the 'contribution' of the first CV (i.e., literature) to the last one (i.e., SDGs). However, there will be DPs in which A connects to B and B to A simultaneously. Therefore, understanding the *type of ties* (as seen in Figure 1) responsible for connecting the *CVs* matters. Usually, the verb that represents the connection among CVs reveals the *type of tie* and the *DP* (e.g., to like, to inform, to help, to advice, to hate, to know, and to act).

Step 3: Data Collection and Selection

In this step, data regarding the CVs previously chosen should be collected. First, considering that scientific data and methods should be replicable, this step should be explained in detail. It is recommended that data be available in the public domain (as public repositories and accessible websites). For instance, in the study of Beck and Ferasso (in press), the 'scientific literature' CV (i.e., articles) was collected from the Scopus database through a Boolean search with complete years of publications limited. Therefore, if the data is derived from a scientific database, the advanced search expression and the day of collection should be written and clearly expressed. As for the 'Sustainable Development Goals' CV, the data used were merely the mission statement of each SDG (Beck & Ferasso, in press), which is publicly available at the United Nations' (2022) website.

Step 4: Data Wrangling

Possible missing values, false-positives and errors should be cleaned from the database. A careful reading of the information of the CVs should be done to assure that the elements address the research purpose. For instance, Beck and Ferasso (in press) read the 45 total elements (i.e. articles) within the 'scientific literature' CV gathered from Scopus and excluded 10 because they were not related to the other CV on 'SDGs'. In other words. Therefore, data wrangling is necessary to assure that the data are related to the CVs, DP, and research purpose.

Qualitative Phase

Step 5: Data Synthesis and Classification

In this step, the researcher starts the *data synthesis* and *classification* following RS by aggregation principles of any CV, not only the literature for which this method was created.

The researcher needs to define the protocol for analyzing the information, and it starts by defining the coding scheme. Considering that there are a wide array of coding scheme possibilities, researchers can code using their knowledge of a subject, a predefined classification, or both. For example, Beck and Ferasso (in press) used the predefined classification where the 17 SDGs were the 17 possible categories. Also, the code form used by Beck & Ferasso (in press) was a table in which literature syntheses according to each of 17 SDGs. The researcher needs to adopt coding decisions, i.e., the procedures to avoid errors when judging ambiguous information and procedures to reduce coder bias and mistakes. Coding decisions can achieve greater accuracy if inspected by another researcher following the same coding scheme.

The retrieved full texts, comprising quantitative, qualitative studies, or both, are thoroughly inspected to identify the *main contributions* to the synthesis. Afterward, a synthesis procedure synthesizes categorical data by reducing the findings/contributions in critical elements for classification and indicating categorical data sources. Therefore, *assessing the quality of information and adherence to the codes* is a critical and intelligible activity of researchers. Furthermore, previously discussed coding decisions strengthen methodological accuracy and reliability.

Departing from the synthesized information, the researcher needs *to integrate the results* by identifying similar syntheses by avoiding duplicate coding and classifying information. For this reason, RS precepts were recommended for CNIA. Then, the researcher relates the integrated results according to the adherence to a specific classification. The coding procedure considers the different characteristics that vary across the syntheses and is built according to the qualitative interpretation of data (inferences). Therefore, the researcher can also apply coding decisions at this point.

The last procedure is the final inspection of the classification according to the adherence of syntheses and classifications and of the inferences. Therefore, another researcher must inspect the reliability of the results by revising all the data synthesis and classification procedures.

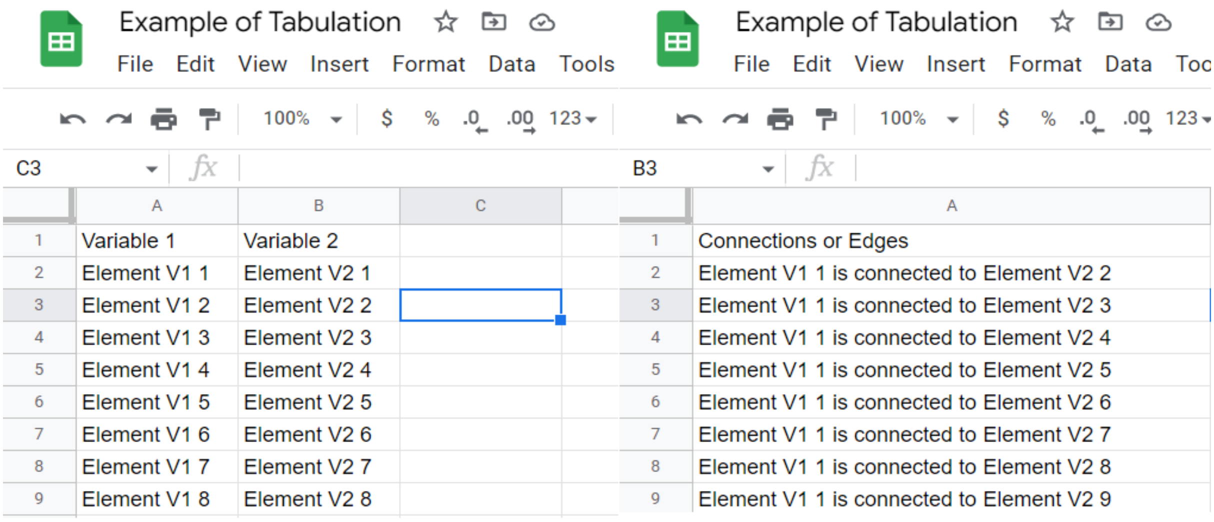
Quantitative Phase

Step 6: Dataset preparation and processing

The data synthesized and classified in step 5 should be transformed into a network by software such as *Gephi* (Bastian et al., 2009) or a programming language such as using the *igraph* package in *R programming language* (Csardi & Nepusz, 2006). It is essential to organize a *new dataset* with the information created in Step 5, thus: (1) two or more variables and their elements should be tabulated, i.e., the elements of each variable are the network nodes; (2) connections between the elements should also be tabulated, i.e., the connections are the network edges; and finally, (3) the data tabulated should be coded/manipulated through the preferred chosen software. Figure 3 illustrates an example of tabulation for the variables with their elements (nodes) and the connections (edges). Therefore, tabulating helps scholars to manipulate data when coding them into a programming language or manipulating them on software.

Figure 3

Example of Tabulation



	A	B	C
1	Variable 1	Variable 2	
2	Element V1 1	Element V2 1	
3	Element V1 2	Element V2 2	
4	Element V1 3	Element V2 3	
5	Element V1 4	Element V2 4	
6	Element V1 5	Element V2 5	
7	Element V1 6	Element V2 6	
8	Element V1 7	Element V2 7	
9	Element V1 8	Element V2 8	

	A
1	Connections or Edges
2	Element V1 1 is connected to Element V2 2
3	Element V1 1 is connected to Element V2 3
4	Element V1 1 is connected to Element V2 4
5	Element V1 1 is connected to Element V2 5
6	Element V1 1 is connected to Element V2 6
7	Element V1 1 is connected to Element V2 7
8	Element V1 1 is connected to Element V2 8
9	Element V1 1 is connected to Element V2 9

Step 7: Network Analysis

After coding, the results for the selected metrics are gathered from the chosen software. Thus, the network can be visualized according to a network layout algorithm, such as *Fruchterman Reingold* and *Yifan Hu* algorithms (Fruchterman & Reingold, 1991; Hu, 2005). Two examples of applications in *Information Networks*: First, Beck and Ferasso (in press, p. 6) chose the *Fruchterman Reingold* algorithm "due to its usability, reliability, and conceptually-intuitive features, better fitting for the research purpose". Second, Beck and Storopoli (2021, p. 6) chose *Yifan Hu* "due to its more realistic demonstration of the dynamism within the network in an efficient and high-quality manner". However, by choosing *Yifan Hu*, nodes can be so far from each other that network visualization can be compromised depending on network characteristics.

Most importantly for CNIA is choosing the metric that best fits to addressing the research purpose. In CNIA, centrality measures in CNIA reveal the most central and peripheral elements within the network. In other words, centrality measures highlight the outstanding elements of the CVs analyzed in CNIA. For this, the *DP* and the *type of tie* should be considered since they are related to the *verb* grounding the *DP* (e.g., to like, to inform, to help, to advice, to hate, to know, to act, etc.). Thus, exploring *centrality measures* in CNIA could reveal the central points of the information network behavior and should be chosen accordingly the research purpose:

First, based on the notion of power as the existing *higher number of connections* to other nodes, *DC* should be chosen to reveal the most *influential* and *prestiged* element (*node*) regarding certain *DP*. In other words, the rationale behind DC is that higher connection to other nodes, higher will be the degree of a node to be *influential* under a *DP* to other nodes.

Second, although often positively correlated to DC (Newman, 2018), *CC* is based on the notion of power as fast accessibility and influenceability to other nodes, which implies understanding which nodes have the *shortest mean distance* to other nodes. CC should be chosen to unveil nodes in strategic locations in the network regarding the *speed* or the *shorter pathways* to be connected to other nodes.

And third, *BC* is based on the notion of power as a node being in the pathway between other nodes, and thus, power is *the ability to connect* others. In other words, it is the ability to create or the essence of having certain *DP*. Therefore, BC should be chosen to unveil nodes in strategic locations in the network regarding the *connectivity role* among nodes. If a node with high BC is removed, the structural distance among the nodes will drastically change and increase.

Step 8: Exploratory Data Analysis

EDA scrutinizes the NA metrics and provides objective and replicable results since it applies the IQR and the whisker limits recommended either by Tukey (1977) or Altman (1991) for the variables studied. In other words, EDA provides more details of properties of the variables. Moreover, EDA provides details of the properties of each variable, such as mean, median, first and third quartiles, mode, and the outliers. These properties are highly recommended to be depicted in boxplots. EDA in CNIA can be performed in three levels of analysis: (1) at the *whole network level* by considering all the clusters, nodes and edges; (2) at the *cluster level* (i.e., CV level), by considering the nodes within a cluster, it is applicable *only* to clusters with more than *5 elements* (see Nuzzo, 2016); and (3) at the *individual level* (i.e., element level), by considering each specific node and its edges, which is *only worth it when analyzing outliers* identified at the whole network and cluster levels. The individual level is particularly made *at the same time* as the whole network or cluster level for identified *outliers*. The bottomline is that EDA reveals the behavior of the elements among the variables as well as if the elements behave normally or abnormally (i.e., for significantly less or more saliency) within a network.

Phase of Meta-Inference

Step 9: Final Analysis

Finally, the researcher performs comparisons between data syntheses and classification (qualitative phase) results and the NA and EDA (quantitative phase) results. Considering the MMR paradigm, CNIA followed the sequential mixed design, where the qualitative phase preceded the quantitative phase (Teddlie & Tashakkori, 2009; Tashakkori et al., 2020). Thus, in each qualitative and quantitative phase, inferences were generated separately. Therefore, the *inferences are integrated into the meta-inference* at this final step.

Inferences represent interpretations and conclusions achieved through a solid understanding process of the collected data. The quality of inferences depends on how the researcher makes sense of results and interpretations when 'connecting the dots.' Therefore, inferences in the MMR paradigm require from the researcher (Tashakkori et al., 2020): (a) high creativity levels of the researcher; (b) insightful intuition; and (c) an ability to reconstruct the aspects of the phenomenon from a 'big picture' perspective.

Meta-inference quality requires carefully inspecting all inferences to guarantee how well they explain the phenomenon altogether (Teddlie & Tashakkori, 2009; Tashakkori et al., 2020). This inspection allows the researcher to build the meta-inference and comprises the integration of qualitative and quantitative inferences. The meta-inference is useful for structuring the analysis and discussion sections.

An exemplary application of the CNIA protocol

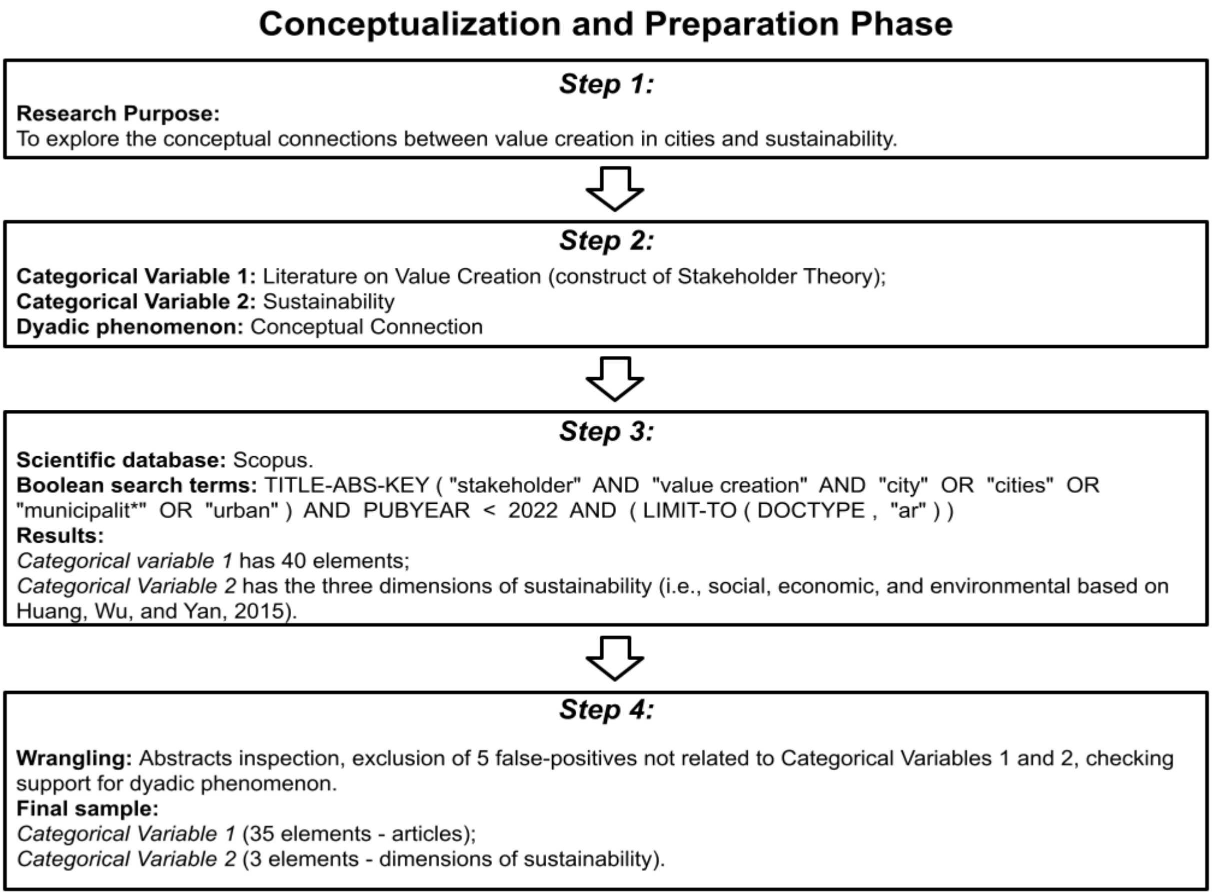
This section illustrates how to use CNIA with an exemplary application, which employed the value creation construct of Stakeholder Theory and the three dimensions of Sustainability.

Conceptualization and Preparation Phase

This example explained how to conduct all the steps of CNIA. Figure 4 illustrates the conceptualization and preparation phase.

Figure 4

Conceptualization and Preparation Phase of Exemplary Application



Analyzing the research purpose makes it possible to identify critical elements for CNIA use, i.e., the *DP* and *CVs*. Since the example considered the literature in one field of study as a CV, the elements (articles) were retrieved from the Scopus database. Additionally, the second CV was predefined, i.e., the three dimensions of sustainability (Huang, Wu & Yan, 2015). After the wrangling procedures, the final sample is prepared for the next phase of CNIA.

Qualitative phase

Step 5 comprises the qualitative phase of the study. In this step, a descriptive synthesis of CV1 is made by simultaneously classifying them and considering the predefined elements of CV2. In other words, the main contributions and findings of the literature on Value Creation of Stakeholder Theory at the city level (composed of 35 articles) were simultaneously synthesized and classified into the three predefined dimensions of sustainability, i.e., social, economic, and environmental.

First, the protocol used for analyzing the information considers the adherence of the elements in CV1 (i.e., articles considering the title, abstract, findings, and conclusions) to CV2 (i.e., dimensions of sustainability).

Second, critical for the synthesis, the *codification process* has a clear coding decision and scheme. About the *coding scheme*, the knowledge of author 1 was the base for coding the articles (i.e., the elements of CV1); and the predefined classification comprised two sustainability dimensions (Huang et al., 2015). As for *coding decisions*, the authors made the decisions consensually, and author 2 revised the coding scheme to avoid dubious information in the articles. Furthermore, after finishing the coding scheme, all procedures were revised by author 2 to avoid bias and mistakes.

In *assessing information quality*, a reducing procedure embodies the main findings and contributions of the codified papers. Thus, the article's corresponding code (i.e., reference) and its synthesis provided the elements used for further classification. The procedure to guarantee the quality of information was how well the synthesis fitted the category, according to the knowledge of author 1. Furthermore, the synthesis and classification procedures were inspected by author 2 to guarantee methodological accuracy and reliability. As result, 32 articles were classified in the Economic dimension classification, 28 in the Social dimension, and 15 in the Environmental dimension.

Integrating results is part of the classification procedure. Accordingly, all the synthesized contributions were inspected to identify how the contributions could be integrated within a given classification. Also, duplicated contributions were integrated into the same synthesis to avoid repetitions. Furthermore, for the classification performed by author 1, the syntheses were inspected to their variance and interpreted qualitatively regarding their content. Lastly, integration was inspected by author 2. The synthesized contributions totalled in 20 central themes (the most salient were smart cities, value creation, innovation, and stakeholder engagement) that were classified according to the three dimensions of sustainability.

In the *final inspection*, the adherence of syntheses and classifications were reinspected by both authors separately, and disagreements were solved after consensus. Thus, the interpretations and conclusions of the syntheses and classifications allow the researcher to produce inferences and information for the next step of the CNIA protocol.

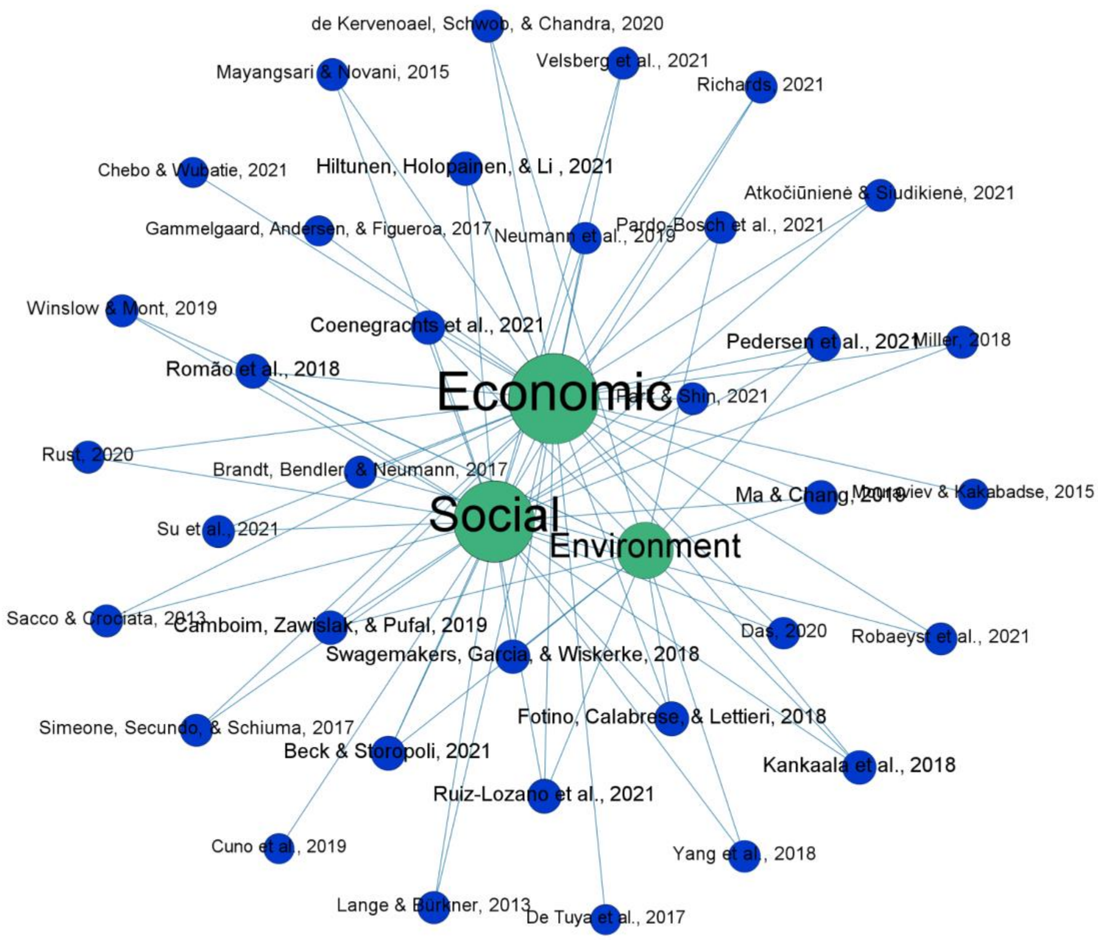
Quantitative phase

In step 6, the results of classifications made in step 5 regarding CV1 and 2 are used as input for a new dataset used in NA and EDA. Finally, the network and its properties were generated through *Gephi*.

In step 7, the NA is performed. It starts by analyzing the results of network properties focusing on the main centrality measures (i.e., DC, CC, and BC). The network visualization is also generated using the algorithm proposed by Fruchterman & Reingold (1991), and in this example, it created a total of 38 nodes and 76 edges.

Figure 5.

Exemplary Network Visualization



The CNIA methodology for the NA considers that two predetermined clusters based on CVs must be formed. Noteworthy is that more than two clusters can be formed according to the needed CVs and research purpose. Then, the NA identifies how the elements and relations are distributed throughout the network, revealing its centralities and visualization. Finally, the researcher must follow the dyadic relations between the two (or more) CVs to interpret the results. Thus, the novel procedure for analyzing the informational network is the *focus on* two (or more) *CVs* exhibited in the network properties, which allows for identifying the *salient* categorical elements in a conceptual/informational DP.

In step 8, EDA is performed to enhance the replicability and objectivity of NA. Importantly, EDA scrutinizes network centrality measures. For example, EDA scrutinized the DC, CC, and BC for revealing the behavior of the two CVs at all network levels.

At the whole network level, results revealed that sustainability's economic and social dimensions were outliers (above the upper data extreme whisker) in the DC, CC, and BC network metrics. Conversely, although the environmental dimension was an outlier only in the DC, it was not an outstanding node in CC and BC measures. Therefore, the concept of 'Value Creation' at the city level has reached only conceptual connections with the economic and social dimensions of 'Sustainability', but the environmental dimension has been underestimated in the literature.
At the cluster level, EDA cannot be applied to 'Sustainability' because it has only three elements. Seven outliers (below the lower data extreme whisker) were identified in 'Literature', none of them connected to the environmental dimension. The same outliers presented the higher mean distance among the nodes. Therefore, there is a weak conceptual connection between these outliers, they possess a peripheral contribution to the DP of the example (i.e., value creation literature and sustainability dimensions).

Therefore, EDA allowed the identification of outliers in the exemplary network by objectifying the NA centrality measurements. Thus, researchers can benefit from EDA at this quantitative phase to strengthen their NA and objectively explain the data obtained from previous network analyses.

Phase of Meta-Inference

In step 9, the inferences were obtained from data syntheses and classification (qualitative phase) and NA and EDA (quantitative phase). Then, these inferences were integrated into a meta-inference.

In the example used for this study, from the syntheses and classifications, the inferences allowed for achieving specific conclusions. In summary, the syntheses revealed that Value Creation is conceptually connected to developing smart sustainable cities, innovation ecosystems, and stakeholder engagement. These elements are critical for sustainability at the city level, but only a few studies adequately address environmental issues.

From the NA and EDA, the main conclusions underlined that Value Creation literature is the most conceptually connected to the economic and social dimensions. However, the challenge is strengthening the connection between Value Creation literature and the underestimated environmental dimension.

The meta-inference from the qualitative and quantitative inferences was built. The main conclusions are: (1) both strands underlined the conceptual connections among the Economic and Social dimensions in the Value Creation literature; and (2) that the environmental dimension needs to be better integrated into the concept of Value creation.

In other words, the qualitative phase provided details of the elements, revealing that smart sustainable cities, innovation ecosystems, and stakeholder engagement play critical roles in the connection between Value Creation and Sustainability, and few studies addressing environmental issues. Furthermore, The quantitative phase reinforced objectively that the environmental sustainability dimension has not been fully integrated into the Value Creation literature.

This meta-inference explained the DP of the used example and was checked by the two authors. Therefore, both inferences were in accordance and enacted complementary, *reinforcing the robustness* of meta-inference. However, the inferences may also present discordant results that the researcher can explain.

Discussion and Conclusion

This research aimed to propose a methodological protocol for analyzing DPs based on two (or more) different CVs by integrating qualitative and quantitative methods. Thus, this research proposed a new methodological protocol, the CNIA, based on the need to explore DPs in-depth and more comprehensively by integrating quantitative and qualitative approaches.

The CNIA comprises synthesis, classification, NA, and EDA because these methods provide more robustness for analyzing DPs. This robustness is achieved thanks to the integrative perspective of qualitative and quantitative methods since their interdependence allows a greater comprehension of the findings. In the CNIA, the syntheses provided the information quality and the main sampled literature contributions that permitted classifications according to predefined categories. These classifications were also relevant for use as input for the NA and EDA. Although both methods presented interconnections, inferences were generated separately and independently. Information NA helps examine complex informational structures and diverse semantics regarding DPs by exploring in-depth CVs. Also, EDA can objectively explain and strengthen the network data performed in network analyses since it can identify outliers and general data distribution measures within the network. The inferences were then integrated to create the meta-inference, which is critical for having a deeper and broader understanding of the DP under study. Therefore, CNIA proved to be useful for any research that explores DPs in any field of study.

The main contribution to MMR is the application for theoretical studies dealing with CVs of DPs. Moreover, although the used example was a review based on literature and predefined classifications, CNIA can also be used by practitioners and policymakers in exploring other CVs (e.g., reports, socioeconomic and demographic data, and any other information sources), and in empirical scholarly research.

Among the limitations, the CNIA considered the syntheses based on human-based interpretations. One limitation could be when a larger quantity of documents must be synthesized. Further investigations could test the efficiency of Natural Language Processing, Machine Learning, or Deep Learning for synthesizing large samples of documents. Considering the NA of CNIA, the categories also can be formed by machine learning or deep learning.

The example used the NA centrality measures; however, CNIA is not limited to these measures. The measures should be chosen according to the research purpose; thus, further investigations may consider new metrics, such as reciprocity, transitivity, homophily, and similarity. Replications of this methodological protocol are recommended in order to increase its improvements.

Acknowledgements and Funding:

No acknowledgements and funding.

Declaration of Interest Statement:
“We, Beck and Ferasso, do not have any conflict of interest”.

Declaration about ethics:
“We, Beck and Ferasso, considered the scientific ethics guidelines”.

References

Aggarwal, C. C. (2017). *Outlier Analysis* (second edition). Cham, Switzerland: Springer Nature.

Agresti, A. (2002). *Categorical Data Analysis*. Hoboken: John Wiley & Sons.

Åkerblad, L., Seppänen-Järvelä, R., & Haapakoski, K. (2021). Integrative strategies in mixed methods research. *Journal of Mixed Methods Research*, *15*(2), 152-170. https://doi.org/10.1177/1558689820957125

Altman, D. G. (1991). *Practical Statistics for Medical Research*. London, England: Chapman and Hall/CRC.

Atkinson, K. M., Koenka, A. C., Sanchez, C. E., Moshontz, H., & Cooper, H. (2015). Reporting standards for literature searches and report inclusion criteria: making research syntheses more transparent and easy to replicate. *Research Synthesis Methods*, *6*(1), 87-95. https://doi.org/10.1002/jrsm.1127

Barbi, A. Q., & Prataviera, G. A. (2019). Nonlinear dependencies on Brazilian equity network from mutual information minimum spanning trees. *Physica A: Statistical Mechanics and its Applications*, *523*, 876-885. https://doi.org/10.1016/j.physa.2019.04.147

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the International AAAI Conference on Web and Social Media*, *3*(1), 361-362. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/13937

Beck, D., & Ferasso, M. (in press). How can Stakeholder Capitalism contribute to achieving the Sustainable Development Goals? A Cross-network Literature Analysis. *Ecological Economics*.

Beck, D., & Storopoli, J. (2021). Cities through the lens of Stakeholder Theory: A literature review. *Cities*, *118*, 103377. https://doi.org/10.1016/j.cities.2021.103377

Beck, D., & Ferasso, M. (2022). Image of Cities as Tool for Urban Governance in Mercosur: Contributions from Urban and City Branding. *Brazilian Journal of Marketing*, *21*(1), 9-28. https://doi.org/10.5585/remark.v21i1.19354

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network Analysis in the Social Sciences. *Science*, *323*(5916), 892-895. https://doi.org/10.1126/science.1165821

Canché, M. S. G. (2022). Network Analysis of Qualitative Data: An Integrative Software Application to Visualize and Assess Similarities in Participants' Qualitative Contributions. *Journal of Mixed Methods Research*, *16*(3), 373-377. https://doi.org/10.1177/15586898211051584

Chalmers, I., Hedges, L. V., & Cooper, H. (2002). A brief history of research synthesis. *Evaluation & the Health Professions*, *25*(1), 12-37. https://doi.org/10.1177/0163278702025001003

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2019). *The handbook of research synthesis and meta-analysis*. New York: Russell Sage Foundation.

Cooper, H. (2015). *Research synthesis and meta-analysis: A step-by-step approach* (Vol. 2). Thousand Oaks: SAGE.

Cordray, D. S., Harris, T. R., & Klein, S. (2009). A research synthesis of the effectiveness, replicability, and generality of the VaNTH challenge-based instructional modules in bioengineering. *Journal of Engineering Education*, *98*(4), 335-348. https://doi.org/10.1002/j.2168-9830.2009.tb01031.x

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*(5), 1–9.

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129-1164. https://doi.org/10.1002/spe.4380211102

Goldsmith, B. E., Semenovich, D., Sowmya, A., & Grgic, G. (2017). Political Competition and the Initiation of International Conflict: A New Perspective on the Institutional Foundations of Democratic Peace. *World Politics*, *69*(3), 493-531. https://doi.org/10.1017/S0043887116000307

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (eighth edition). Andover: Cengage Learning, EMEA.

Hedges, L. V., & Cooper, H. (2009). Research synthesis as a scientific process. In H. Cooper, L. V. Hedges, & Valentine J. C. (Eds.). *The handbook of research synthesis and meta-analysis*, New York: Russel Sage Foundation, 3-18.

Heyvaert, M., Maes, B., & Onghena, P. (2013). Mixed methods research synthesis: definition, framework, and potential. *Quality & Quantity*, *47*(2), 659-676. https://doi.org/10.1007/s11135-011-9538-6

Hu, Y. (2005). Efficient, High-Quality Force-Directed Graph Drawing. *The Mathematica Journal*, *10*(1), 37–71.

Huang, L., Wu, J., & Yan, L. (2015). Defining and measuring urban sustainability: A review of indicators. *Landscape Ecology*, *30*(7), 1175-1193. https://doi.org/10.1007/s10980-015-0208-2

Jacobs, D. B., & Cramer, L. A. (2017). Applying information network analysis to fire-prone landscapes: implications for community resilience. *Ecology and Society*, *22*(1), 52. https://doi.org/10.5751/ES-09119-220152

Kim, J., Yammarino, F. J., Dionne, S. D., Eckardt, R., Cheong, M., Tsai, C. Y., ... & Park, J. W. (2020). State-of-the-science review of leader-follower dyads research. *The Leadership Quarterly*, *31*(1), 101306. https://doi.org/10.1016/j.leaqua.2019.101306

Korsgaard, M. A., Brower, H. H., & Lester, S. W. (2015). It isn't always mutual: A critical review of dyadic trust. *Journal of Management*, *41*(1), 47-70. https://doi.org/10.1177/0149206314547521

Krasikova, D. V., & LeBreton, J. M. (2012). Just the two of us: Misalignment of theory and methods in examining dyadic phenomena. *Journal of Applied Psychology*, *97*(4), 739. https://doi.org/10.1037/a0027962

Laudan, L. (1981). *Science and hypothesis: Historical essays on scientific methodology*. Netherlands: Springer Science. https://doi.org/10.1007/978-94-015-7288-0

Lyons, K. S., & Lee, C. S. (2018). The theory of dyadic illness management. *Journal of Family Nursing*, *24*(1), 8-28. https://doi.org/10.1177/1074840717745669

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474-496. https://doi.org/10.1002/tea.20347

Newman, M. (2018). *Networks* (second edition). Oxford: Oxford University Press.

Nooraie, R. Y., Sale, J. E., Marin, A., & Ross, L. E. (2020). Social network analysis: An example of fusion between quantitative and qualitative methods. *Journal of Mixed Methods Research*, *14*(1), 110-124. https://doi.org/10.1177/1558689818804060

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*(3), 417-528. https://doi.org/10.1111/0023-8333.00136

Nuzzo, R. L. (2016). The box plots alternative for visualizing quantitative data. *PM&R*, *8*(3), 268-272. http://dx.doi.org/10.1016/j.pmrj.2016.02.001

Orwin, R. G. (1994). Evaluating coding decisions. In H. Cooper & L. V. Hedges (Eds.). *The Handbook of Research Synthesis.* New York: Russell Sage Foundation, 139-162.

Sandelowski, M., Voils, C. I., Leeman, J., & Crandell, J. L. (2012). Mapping the mixed methods–mixed research synthesis terrain. *Journal of Mixed Methods Research*, *6*(4), 317-331. https://doi.org/10.1177/1558689811427913

Schmitt, J. B., Rieger, D., Rutkowski, O., & Ernst, J. (2018). Counter-messages as prevention or promotion of extremism?! the potential role of youtube: Recommendation algorithms. *Journal of Communication*, *68*(4), 758-779. https://doi.org/10.1093/joc/jqy029

Shi, C., Li, Y., Yu, P. S., & Wu, B. (2016). Constrained-meta-path-based ranking in heterogeneous information network. *Knowledge and Information Systems*, *49*(2), 719-747. https://doi.org/10.1007/s10115-016-0916-1

Shi, C., Li, Y., Zhang, J., Sun, Y., & Yu, P. S. (2017). A Survey of Heterogeneous Information Network Analysis. *IEEE Transactions on Knowledge and Data Engineering*, *29*(1), 17-37. https://doi.org/10.1109/TKDE.2016.2598561

Spitzer, M., Wildenhain, J., Rappsilber, J., & Tyers, M. (2014). BoxPlotR: a web tool for generation of box plots. *Nature Methods*, *11*(2), 121-122.https://doi.org/10.1038/nmeth.2811

Stock, W. A. (1994). Systematic coding for research synthesis. In H. Cooper & L. V. Hedges (Eds.). *The Handbook of Research Synthesis.* New York: Russell Sage Foundation, 125-138.

Sun, Y., Han, J., Yan, X., & Yu, P. S. (2012). Mining knowledge from interconnected data: A heterogeneous information network analysis approach *Proceedings of the VLDB Endowment*, *5*(12), 2022-2023. https://doi.org/10.14778/2367502.2367566

Suri, H. (2011). Purposeful sampling in qualitative research synthesis. *Qualitative Research Journal*, *11*(2), 63-75. https://doi.org/10.3316/QRJ1102063

Tashakkori, A., Johnson, R. B., & Teddlie, C. (2020). *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences* (2nd edition). Thousand Oaks: Sage Publications.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Thousand

Oaks: Sage.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, Massachusetts: Addison-Wesley.

United Nations (2022). *The 17 Goals | Sustainable Development*. Retrieved from: https://sdgs.un.org/goals

Voils, C. I., Sandelowski, M., Barroso, J., & Hasselblad, V. (2008). Making sense of qualitative and quantitative findings in mixed research synthesis studies. *Field Methods*, *20*(1), 3-25. https://doi.org/10.1177/1525822X07307463

Wang, J., Chen, C., Li, H. F., Jiang, X. L., & Zhang, L. (2016). Investigating key genes associated with ovarian cancer by integrating affinity propagation clustering and mutual information network analysis. *European Review for Medical and Pharmacological Sciences*, *20*(12), 2532-40.

Yu, P. S., Han, J., & Faloutsos, C. (2010). *Link mining: Models, Algorithms, and Applications*. New York: Springer Nature.

Xie, Y., Yu, B., Lv, S., Zhang, C., Wang, G., & Gong, M. (2021). A survey on heterogeneous network representation learning. *Pattern Recognition*, *116*, 107936. https://doi.org/10.1016/j.patcog.2021.107936