



UvA-DARE (Digital Academic Repository)

Network Models to Organize a Dispersed Literature: The Case of Misunderstanding Analysis of Covariance

Derks, K.; Burger, J.; van Doorn, J.; Kossakowski, J.J.; Matzke, D.; Atticciati, L.; Beitner, J.; Benzesin, V.; de Bruijn, A.L.; Cohen, T.R.H.; Cordesius, E.P.A.; van Dekken, M.; Delvendahl, N.; Dobbelaar, S.; Groenendijk, E.R.; Hermans, M.E.; Hiekkaranta, A.P.; Hoekstra, R.H.A.; Hoffmann, A.M.; Hogenboom, S.A.M.; Kahyeci, S.; Karaban, I.J.; Kevenaer, S.; te Koppele, J.L.; Kramer, A-W.; Kroon, E.; Kucharský, Š.; Lieuw-On, R.; Lunansky, G.; Matzen, T.P.; Meijer, A.; Nieper, A.; de Nooij, L.; Poelstra, L.; van der Putten, W.J.; Sarafoglou, A.; Schaaf, J.V.; van de Schraaf, S.A.J.; van Schuppen, S.; Schutte, M.H.M.; Seibold, M.; Slagter, S.K.; Snoek, A.C.; Stracke, S.; Tamimy, Z.; Timmers, B.; Tran, H.; Uduwa-Vidanalage, E.S.; Vergeer, L.; Vossoughi, L.; Yücel, D.E.; Wagenmakers, E.-J.

DOI

[10.5334/jeps.458](https://doi.org/10.5334/jeps.458)

Publication date

2018

Document Version

Final published version

Published in

Journal of European Psychology Students

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Derks, K., Burger, J., van Doorn, J., Kossakowski, J. J., Matzke, D., Atticciati, L., Beitner, J., Benzesin, V., de Bruijn, A. L., Cohen, T. R. H., Cordesius, E. P. A., van Dekken, M., Delvendahl, N., Dobbelaar, S., Groenendijk, E. R., Hermans, M. E., Hiekkaranta, A. P., Hoekstra, R. H. A., Hoffmann, A. M., ... Wagenmakers, E.-J. (2018). Network Models to Organize a Dispersed Literature: The Case of Misunderstanding Analysis of Covariance. *Journal of European Psychology Students*, 9, 48-57. <https://doi.org/10.5334/jeps.458>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

RESEARCH ARTICLE

Network Models to Organize a Dispersed Literature: The Case of Misunderstanding Analysis of Covariance

Koen Derks, Julian Burger, Johnny van Doorn, Jolanda J. Kossakowski, Dora Matzke, Ludovica Atticciati, Julia Beitner, Viket Benzesin, Anne L. de Bruijn, Tara R. H. Cohen, Elisa P. A. Cordesius, Marit van Dekken, Nora Delvendahl, Simone Dobbelaar, Eva R. Groenendijk, Merel E. Hermans, Anu P. Hiekkaranta, Ria H. A. Hoekstra, Agnes M. Hoffmann, Sally A. M. Hogenboom, Sercan Kahveci, Irina J. Karaban, Sofieke T. Kevenaer, Jurriaan L. te Koppele, Anne-wil Kramer, Emese Kroon, Šimon Kucharský, Ricardo Lieuw-On, Gaby Lunansky, Timo P. Matzen, Annemarie Meijer, Annika Nieper, Laura de Nooij, Leonie Poelstra, Wikke J. van der Putten, Alexandra Sarafoglou, Jessica V. Schaaf, Sara A. J. van de Schraaf, Steven van Schuppen, Manon H. M. Schutte, Mitja Seibold, Scarlett K. Slagter, Aishah C. Snoek, Selina Stracke, Zenab Tamimy, Bram Timmers, Han Tran, Elizabeth S. Uduwa-Vidanalage, Laura Vergeer, Linos Vossoughi, Dilan E. Yücel and Eric-Jan Wagenmakers

We outline a network method to synthesize a literature overview from search results obtained by multiple team members. Several network statistics are used to create a single representativeness ranking. We illustrate the method with the dispersed literature on a common misinterpretation of analysis of covariance (ANCOVA). The network method yields a top ten list of the most relevant articles that students and researchers can take as a point of departure for a more detailed study on this topic. The proposed methodology is implemented in Shiny, an open-source R package.

Keywords: Literature review; ANCOVA; Networks; Combining ranks; Data augmentation

Introduction

Analysis of covariance (ANCOVA) remains a widely misunderstood approach for dealing with group differences on potential covariates (Miller & Chapman, 2001). This misunderstanding of the ANCOVA has a long history and its discussion is dispersed across fields and journals, making it difficult to obtain a systematic overview. Here we present a network method to organize the results of a literature search conducted by 44 Master's students as part of the 2016 University of Amsterdam course "Good Research Practices".

The ANCOVA Pitfall

Dora wants to assess whether, in her own university, men earn more than women. She has access to the salaries of a subset of researchers, and, as expected, men earn significantly more than women ($p < .005$). But wait! The

men in her sample are also older than the women, and this confounds the results: perhaps the salary difference is due to age rather than gender. To address this confound and "control for" age, Dora includes age as a covariate in an ANCOVA. This procedure is tempting but statistically problematic. The ANCOVA is easier to interpret correctly when age influences salary but does *not* differ across the groups.

As explained in Miller and Chapman (2001; but see chapter 10 in Judd, McClelland, & Ryan, 2011, and Field, 2013, pp. 484–486), when groups differ on a covariate (e.g., age), removing the variance associated with the covariate also removes the shared variance associated with the group (e.g., gender). As a result, the grouping variable loses some of its representativeness. This occurs mostly when groups are pre-existing and are not obtained by random assignment (Jamieson, 2004). As an example, assume one has access to the height of several mountain peaks in the Himalayas and the Pyrenees (Cohen & Cohen, 1983). One may test whether the mountain ranges differ in height and it may be tempting to include air pressure as a covariate; after all, air pressure differs across the

mountain ranges, confounding the results. However, air pressure is intimately related to elevation, and removing the variance in elevation associated with air pressure removes virtually all of the variance in elevation associated with the mountain range. The outcomes of the ANCOVA may suggest that, controlling for air pressure, the mountain ranges are about equally high. In other words, the result of an ANCOVA with confounded covariates is problematic: it invites the false interpretation that an inherent confound can be removed by purely statistical means, by “controlling” for it, whereas the correct interpretation is: “Having accounted for the confounded covariate, what is the added value of including the treatment effect?”

Ways to Organize a Literature

As detailed below, 44 students selected articles from the pertinent literature (i.e., work on the misunderstanding of ANCOVA), and rated their relevance (i.e., the importance and informativeness of the article with respect to the problem under consideration). One can organize and synthesize the resulting data in different ways. For instance, one could focus on the frequency with which the students reported an article. The disadvantage is that hidden treasures—valuable articles found by only a few students—are overlooked. Another method is to consider the average relevance rating that an article receives; the disadvantage here is that this method ignores the wisdom of the group, as an article that has been found by a single user and rated “10” may falsely appear as highly important. In addition, neither of these two methods quantifies the possible association between different articles.

As an alternative way to analyze and visualize the outcomes of the literature search, we outline two network models. A network model yields a flexible representation of the importance of objects and their relationships. Its distinguishing feature is that the schema can be viewed as a graph in which object types are nodes and relationships between the objects are edges. For instance, in a network of depression, the nodes correspond to the symptoms (e.g., sleep loss, fatigue, loss of appetite) and edges between the symptoms quantify the strength of association, such that a pronounced association between sleep loss and fatigue will be reflected by a prominent edge between the nodes that correspond to those symptoms (e.g., Borsboom & Cramer, 2013). In addition, one can analyze properties of individual nodes in the graph, such as their importance (i.e., how strongly a particular node is connected to all the other nodes in a network; Costantini et al., 2015). The idea of organizing literature in networks is not new (Wijngaert, Bouwman, & Contractor, 2014). In the Wijngaert et al. approach, nodes are concepts that occur in articles, and edges represent a relation that was stated as a hypothetical relation between concepts. For instance, awareness of government technology and the actual use of this technology are nodes in the Wijngaert network. They are connected by the causal hypothesis that awareness affects actual use. These networks offer an insightful overview of concepts and therefore summarize the conceptual relations across many different articles. The network approach proposed in this paper differs from the network model in

Wijngaert et al.: we observe articles instead of concepts and therefore the nodes represent individual articles. Using network models, we hope to identify papers that are most relevant in a certain field of interest.

We first used the *qgraph* package in R (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012) to create the networks, with articles represented by nodes and relationships between articles by edges. Next, we computed several centrality measures for each article, and then used a Bayesian rank-based inference methodology to arrive at an overall rating of each article's importance.

Method

The literature search was conducted by 44 students in the 2016 Research Master's course “Good Research Practices” at the Department of Psychology of the University of Amsterdam. Each student collected 40 articles published on the ANCOVA problem: 20 articles published prior to the seed article by Miller and Chapman (2001), and 20 articles published after. Each student rated the relevance of each of their chosen articles on a scale from one to ten. After students selected the articles and rated their relevance, we created two networks to visualize the results, showing co-occurrences and citations as edges and articles as nodes.

Note that the literature search departed from knowledge of the Miller and Chapman (2001) seed article; this high-impact article clearly outlines the problem and its history. It is likely that the presence of this seed article (together with the articles in its reference list) partly guided the literature search, thereby creating results that are substantially more homogeneous than would otherwise have been the case.

Network Models

For concreteness, consider the hypothetical networks shown in **Figure 1**, based on mock data shown in **Table 1**. Here, a literature search has been performed by five participants, who in total found four different articles. The table shows for each individual rater which articles he or she found. We will use this toy example to clarify key concepts in two different networks: a co-occurrence network and a citation network. Both network methods have the advantage that they do not just describe the data, they also visualize important relations between articles.

Approach I: Co-occurrence network. A co-occurrence network is a network of the process of the literature search, and more specifically of the co-occurrence of articles within raters. Nodes represent articles and edges represent the number of times the articles were found together by raters. All of the edges in the network are undirected, as there are no causal connections between article pairs. One advantage of this network method is that it is not constrained to a timeline, as articles of all years can be found together. One disadvantage is that hidden treasures might not be considered relevant, as centrality in this network is mostly determined by how often articles were found.

To quantify how central an article is, the co-occurrence network method uses several centrality measures. These centrality measures can be seen as indicators of the

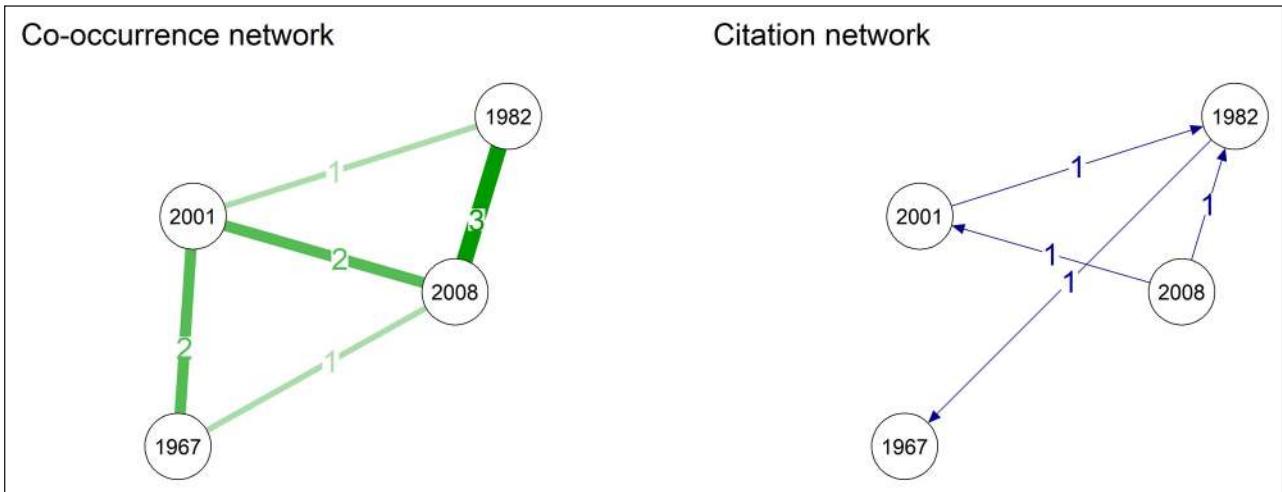


Figure 1: Two different networks. On the left, a co-occurrence network based on the mock data in Table 1 is visualized. On the right, a citation network based on the mock data in Table 1 is visualized.

Table 1: Mock data from five raters who each reported two or three relevant articles in random order. The year uniquely identifies an article. (e.g., the article “1982” is reported by raters 1, 2, and 5).

Rater	Article 1	Article 2	Article 3
1	1982	2008	
2	2001	2008	1982
3	1967	2008	2001
4	1967	2001	
5	1982	2008	

importance of that node (Costantini et al., 2015) and consist of a number of indices, depending on whether the network has directed connections or undirected connections. For an undirected network, the centrality measures consist of *betweenness*, *closeness*, and *strength*. The *betweenness-index* deals with the question of how well one specific node connects other nodes. The papers that score high in betweenness therefore connect many different papers with one another, by virtue of being chosen by multiple raters. In our example, we see that there are several nodes that are connected to other nodes. **Table 2** shows the centrality measures for our example.

The node from 2008 has the highest betweenness, because it connects every article with each other and has paths that are stronger. The *closeness-index* represents how easy it is to reach all other nodes from one specific node. If it takes only a few steps to get from one article to all other articles, we observe a high closeness. For our example, both the nodes 2001 and 2008 have the highest closeness, because they have the most connections to the other nodes. The *strength-index* deals with the question how well one node is connected with all other nodes. It is a weighted measure and represents the sum of the edge weights going in and out of the node. For our example, we observe that the node 2008 has the highest strength, because it is most strongly connected to other nodes (i.e., it has the highest edge strengths).

Table 2: Centrality measures for the co-occurrence network on the mock data.

Article	Betweenness	Closeness	Strength
1967	-0.72	-1.08	-1.16
1982	-0.72	-0.61	-0.38
2001	0.06	0.84	0.38
2008	1.39	0.84	1.16

Approach II: Citation network. A citation network visualizes how articles refer to one another. In this network structure, every article is again represented as a node. Directed edges between two nodes represent a connection between two articles in terms of a citation structure. An arrow from article X to article Y means that article X cites article Y, whereas the opposite means that article Y cites article X. To quantify how central a node is in a citation network, we have to use different centrality measures than in the co-occurrence network.

For a directed network, the centrality measures consist of *betweenness*, *closeness*, *in-degree*, and *out-degree*. Here, the *strength-index* is split up into the *in-degree* and *out-degree* index. These indices are unweighted indices and respectively represent how often an article gets a connection and how often an article gives a connection to other articles. For an overview of centrality measures in networks see for instance Friedkin (1991), White and Borgatti (1994), and Stephenson and Zelen (1989). The centrality measures for the citation network are shown in **Table 3**.

The *in-degree* centrality index shows which articles are referred to the most, whereas the *betweenness-index* shows which articles connect to the ideas of other articles the most (Costantini et al., 2015). For our example, the node from 1982 has been referred to the most, and thus has the highest in-degree index. It also has the highest betweenness index, because it acts as a path to other nodes the most. The node 2008 has the highest *out-degree* index, because it cites other articles the most. These centrality indices give us an indication of relevance,

and thus may help us to create a more efficient reading list. Reading a highly central article is efficient, because such an article is likely to adopt ideas that are addressed in many other papers on the same issue and thus might make other papers redundant. The main advantage of a citation network is that it visualizes impact, that is, how many articles refer to that specific article. The main disadvantage of the citation network is a timeline restriction: earlier articles cannot cite later articles.

A citation network is created as follows. First, we create an *edge list* with the first column containing the source articles, which are the articles that cite others. The second column contains, for each source article, all of the other articles that the source article could possibly refer to. These include all of the remaining articles that have been collected. The third column contains a binary indicator variable: 1 for a reference, 0 for no reference. Before being able to create a citation network we have to prune the edge list so that it contains only the most relevant articles. To this aim we first proceeded to include the 15 articles with the highest average relevance rating and a minimum of five student nominations. To include possible hidden treasures in the set, we then added the five articles with the highest average relevance rating and a maximum of four student nominations.

We wrote an R script and an R Shiny web application to create adjacency matrices for the networks, plot the networks, apply the Bayesian rank-based method, and implement this method for one’s own literature synthesis. These can be found in the online appendix (osf.io/bxz6c/), together with all of the data.

Combining Ranks

In both networks, we can rank articles according to their corresponding network-specific importance measures (i.e., the centrality indices). In order to arrive at a single importance ranking, the importance measures need to be aggregated across the different centrality indices. This aggregation can be accomplished by introducing a latent, normally distributed variable of importance whose values are constrained by the observed ordinal information. In statistics, this technique is known as data augmentation, and the latent variable can be estimated through Gibbs sampling. Thus, for each ordinal data point (i.e., rank), its corresponding latent value of importance is estimated by means of a posterior distribution, that is, a representation of uncertainty across the different importance values. In the past, this technique has been used to estimate the polychoric correlation coefficient

Table 3: Centrality measures for the citation network on the mock data.

Article	Betweenness	In-degree	Out-degree
1967	-0.5	0	-1.22
1982	1.5	1.22	0
2001	-0.5	0	0
2008	-0.5	-1.22	1.22

(Albert, 1992) and Kendall’s tau (van Doorn, Ly, Marsman, & Wagenmakers, 2017; for an overview of introductory materials on Bayesian inference see, for instance, Etz et al., 2017). In the current context, however, the estimates reflect the latent importance of each article. Applying this procedure to each of the three centrality indices therefore gives, for each article, three posterior distributions of its latent importance. In order to aggregate these estimates, we computed the average of the three posterior medians. Doing so for all articles results in a new ranking of the articles that combines each of the three centrality measures of importance. We will refer to this new ranking as the aggregated importance ranking.

For illustration, we apply the network method to the subset of 233 articles in the literature before 2001. Results for the literature after 2001 are presented in online Appendix A (osf.io/bxz6c/).

Results

Descriptive Method

To illustrate how one could create a top-ten list of relevant articles without the benefit of a network model, we show the results of a descriptive method to rank articles on relevance. Based on ad-hoc cut-offs, we included articles only when reported by a minimum of 10 raters and with a mean relevance grade of at least 8. The result is shown in **Table 4**.

Co-occurrence Network

We estimated the co-occurrence network using all articles found by the 44 raters before 2001. The network is displayed in **Figure 2**; centrality measures were computed by the *qgraph* package in R (Epskamp, Cramer, Waldorp, Schmittmann & Borsboom, 2012). The centrality measures for the articles are displayed in **Figure 3**.

We then computed Kendall’s tau, a rank-based correlation coefficient, for the correlations between the article’s relevance ratings and the various centrality measures (van Doorn et al., 2017). As shown in **Table 5**, the relevance grades—which have no direct impact on the connections in the co-occurrence network—correlate positively with the centrality measures of the articles. Although the

Table 4: Articles found by at least 10 raters and with a minimum mean relevance grade of 8.

Article	Found by	Mean relevance grade (SD)
Evans & Anastasio (1968)	42	8.8 (1.2)
Lord (1967)	39	8.6 (1.3)
Elashoff (1969)	36	8.4 (1.2)
Overall & Woodward (1997)	32	8.3 (0.9)
Lord (1969)	26	8.4 (1.3)
Adams et al. (1985)	21	8.1 (1.2)
Glass et al. (1972)	13	8.1 (1.2)
Storandt & Hudson (1975)	13	8.2 (1.1)

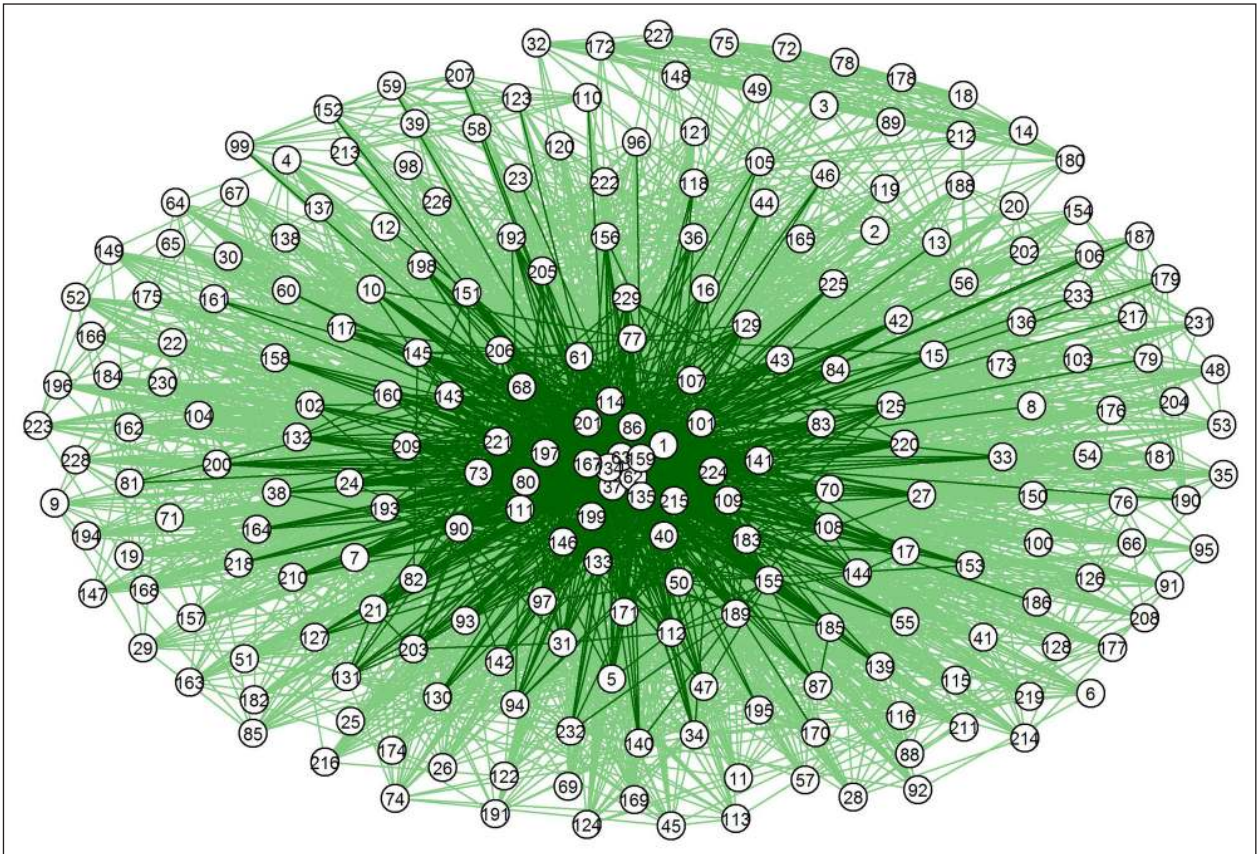


Figure 2: The co-occurrence network applied to the search process for the ANCOVA-pitfall literature before 2001. Each node represents a reported article. Undirected edges represent the number of times the two articles were reported by the same rater. Thicker and saturated edges represent stronger connections.

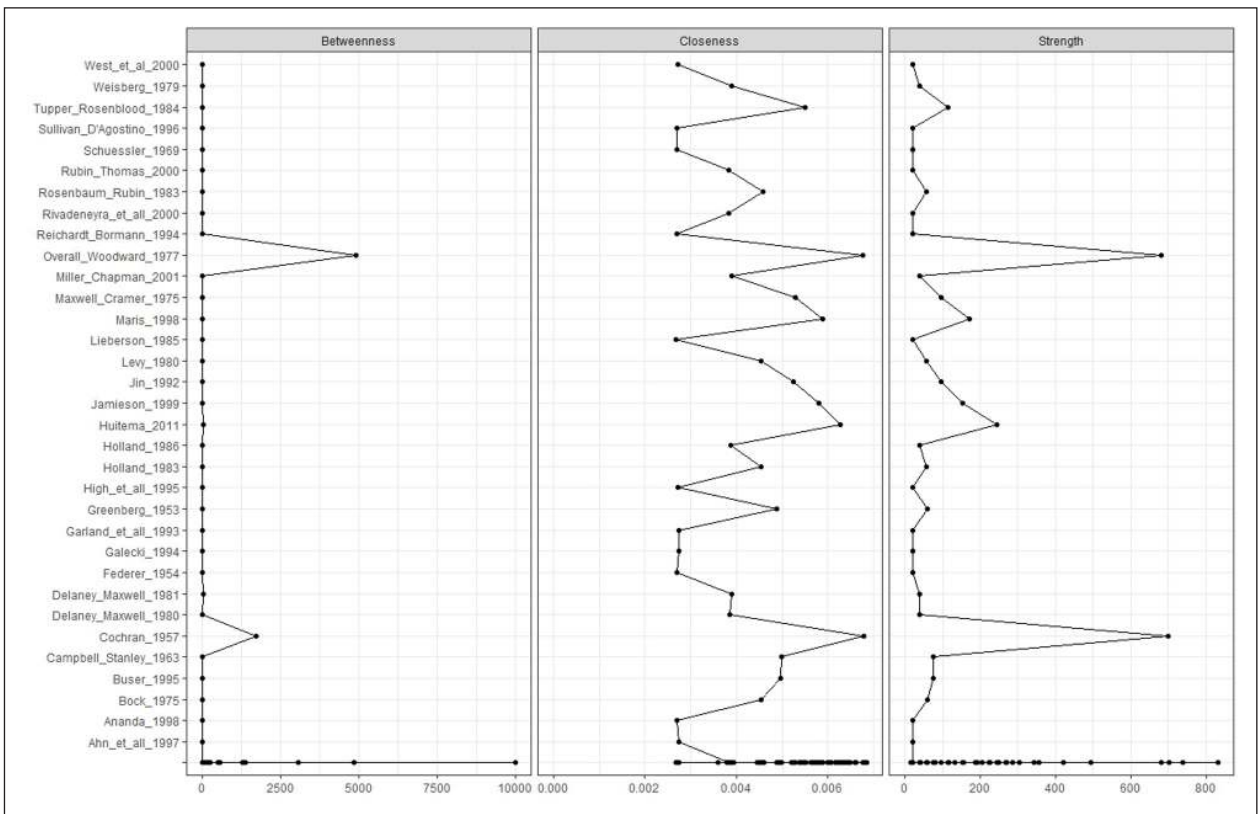


Figure 3: Centrality plot containing all centrality indices for the single nodes of the co-occurrence network. For readability, only 44 of the 244 articles are shown.

correlations are low, this does suggest that the network captures some of the unseen information that is contained in the relevance grades.

The aggregated importance ranks of the co-occurrence network were then used to construct a list of the top ten most relevant articles. This list is shown in **Table 6**.

As shown in **Table 7**, the aggregated importance ranks (i.e., the final network centrality ranks based on the aggregation of the three centrality measures through data augmentation) correlate positively with ranks obtained from the number of raters, from the mean relevance grade, and the average of the two. The co-occurrence network method captures information similar to information obtained when ranking based on the number of raters; however, the network centrality measures provide an additional important characteristic of the articles, namely their co-occurrence with relevant articles.

Table 5: Kendall correlations and 95% credible intervals for the correlations between the mean relevance grades and centrality measures of the articles used in the co-occurrence network.

	τ	95% credible interval
Relevance grade – Betweenness	0.255	[0.164, 0.337]
Relevance grade – Closeness	0.224	[0.134, 0.307]
Relevance grade – Strength	0.258	[0.166, 0.341]

Table 6: Top ten articles in the entire search before 2001. The ranks are the final network ranks obtained from the co-occurrence network.

Article	Title	Rank
Evans & Anastasio (1968)	Misuse of analysis of covariance when treatment effect and covariate are confounded.	1
Lord (1967)	A paradox in the interpretation of group comparisons.	2
Overall & Woodward (1977)	Nonrandom assignment and the analysis of covariance.	3
Cochran (1957)	Analysis of covariance: Its nature and uses.	4
Elashoff (1969)	Analysis of covariance: A delicate instrument.	5
Porter & Raudenbush (1987)	Analysis of covariance: Its model and use in psychological research.	6
Adams et al. (1985)	Analysis of covariance as a remedy for demographic mismatch of research subject groups: Some sobering simulations.	7
Lord (1969)	Statistical adjustments when comparing preexisting groups.	8
Wainer (1991)	Adjusting for differential base rates: Lord’s paradox again.	9
Keselman et al. (1998)	Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analysis.	10

Table 7: Kendall correlations and 95% credible intervals computed between the final rank based on the co-occurrence network and ranks based on mean relevance grade and number of raters.

	τ	95% credible interval
A: Rank by number of raters – Final rank	0.789	[0.687, 0.858]
B: Rank by relevance grade – Final rank	0.241	[0.150, 0.325]
Rank by average of A and B – Final rank	0.554	[0.459, 0.631]

Citation Network

For the construction of the citation network we selected a subset of 20 key articles from all raters’ nominations using the following criteria: (a) select 15 articles with the highest average relevance grade based on a minimum of five nominations; (b) select five articles with the highest average relevance grade based on a maximum of four nominations; and (c) in case of ties, include all articles with the same relevance grade.

As there were several ties in the relevance grades, the selection criteria yielded 26 articles in total. For the 15 best-graded articles with a minimum of five nominations, we observed a cut-off inclusion grade of 7.86. Articles with tied scores were added to the set, resulting in the selection of 17 articles. For the 5 best-graded articles with a maximum of four nominations we observed a cut-off inclusion grade of 9.00. Articles with tied scores were also added to the set, resulting in the selection of nine articles. The resulting network can be seen in **Figure 4**. Online appendix B (osf.io/bxz6c/) contains the same network but includes Miller and Chapman (2001) to visualize the impact of their paper.

The citation network is not fully connected, because some articles do not cite any of the other articles, and do not receive any citations from those articles. Articles (i.e., nodes) with many incoming arrows are cited relatively often. The citation network can be analyzed using the same technique we applied to the co-occurrence network. **Figure 5** shows the centrality measures for the citation

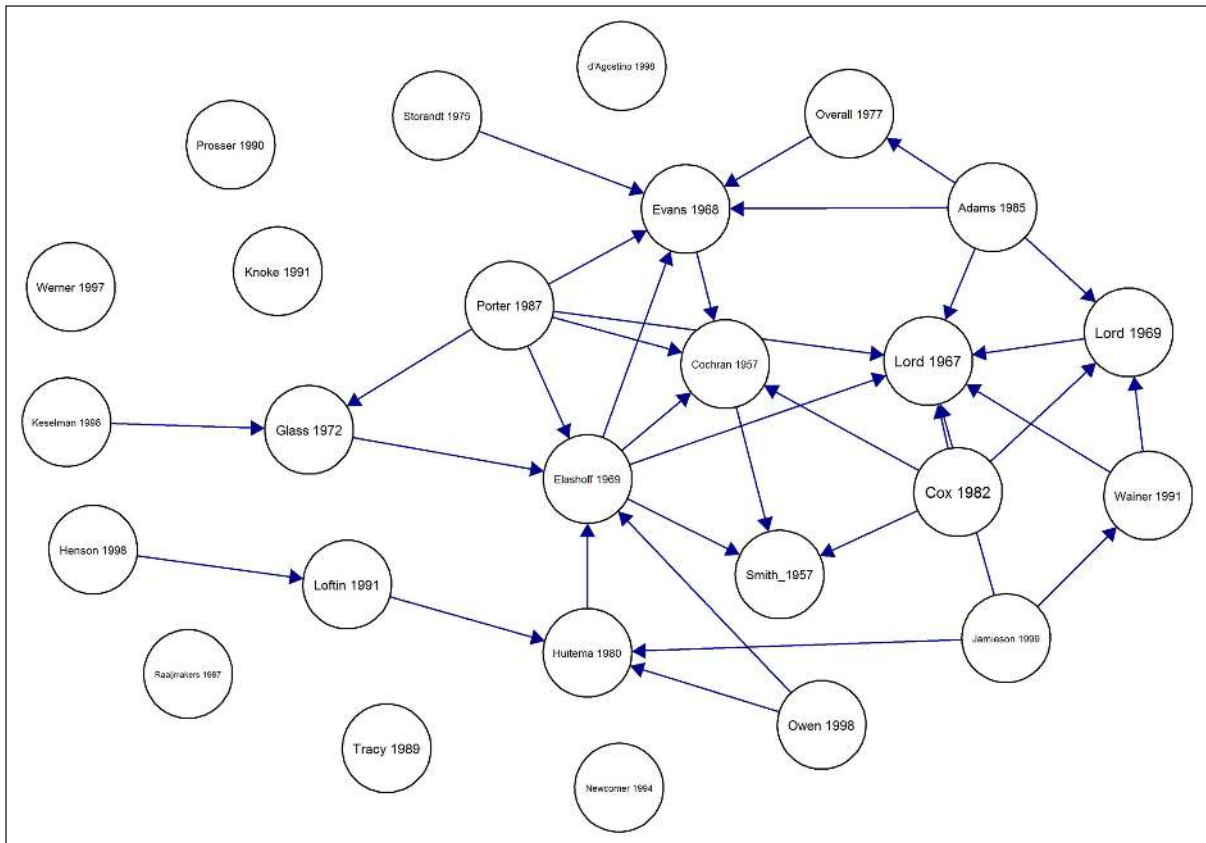


Figure 4: The citation network applied to the search process for the ANCOVA-pitfall literature before 2001. Each node represents a reported article. Directed edges represent whether or not an article cited another article. Connected nodes represent a source-target relation. The source is the paper that cites the target, for example the arrow from Cox (1982) to Lord (1967) means that Cox cited Lord.

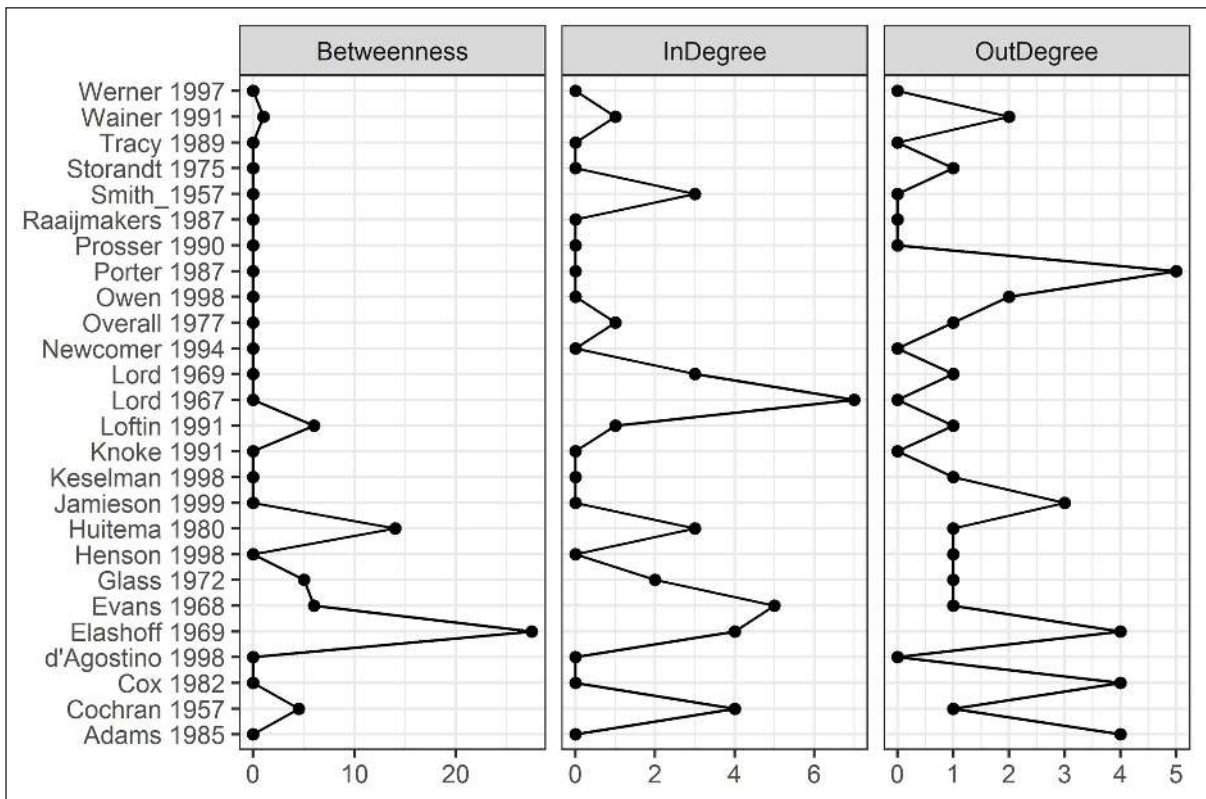


Figure 5: Centrality plot containing all centrality indices for the single nodes. Note that closeness requires a fully connected network and is therefore omitted.

network. Closeness is omitted from this plot because it requires a fully connected network.

As before, we computed correlations between the centrality indices and the mean relevance grades. As can be seen in **Table 8**, the mean relevance grades do not appear to correlate with the centrality measures of the articles. This may be explained by the fact that more recent articles can be highly relevant (e.g., they can review and synthesize the earlier literature) but these articles cannot be cited by articles that were published earlier. In other words, the relevance grades are not subject to the temporal restriction that governs the citation patterns.

The aggregated importance ranks of the citation network were then used to construct a top ten list of relevant articles. This list is shown in **Table 9**.

Table 8: Kendall correlations and 95% credible intervals computed between the mean relevance grades and centrality measures of the articles used in the citation network.

	τ	95% credible interval
Relevance grade – Betweenness	-0.090	[-0.333, 0.168]
Relevance grade – In-degree	0.012	[-0.240, 0.261]
Relevance grade – Out-degree	-0.192	[-0.423, 0.078]

As shown in **Table 10**, the aggregated importance ranks (i.e., the final network centrality ranks based on the aggregation of the three centrality measures through data augmentation) do not correlate with ranks obtained from the number of raters, from the mean relevance grade, and the average of the two. This suggests that a citation network alone is not sufficient for assessing article relevance.

Concluding Comments

We used two networks to organize a dispersed literature. Compared to the descriptive method (i.e., a list of the average relevance grade and the number of times an article is reported), the networks allow one to discover relationships between articles, and discover hidden treasures. Nevertheless, the value of the network approach over the descriptive method warrants future empirical scrutiny.

For this particular project we started with one particular seed paper (i.e., Miller & Chapman, 2001) which clearly influenced the results. More heterogenous outcomes can be expected when students are not given a seed paper to start with.

To facilitate the use of the network method we recommend an additional resource over and above the *qgraph* R package. We have developed an R shiny web application (<https://koenderks.shinyapps.io/LiteratureNetworks/>) that implements functions to estimate a co-occurrence

Table 9: Top ten articles in the entire search before 2001. The ranks are the final network ranks obtained from the citation network.

Article	Title	Rank
Elashoff (1969)	Analysis of covariance: A delicate instrument.	1
Huitema (1980)	The analysis of covariance and alternatives.	2
Evans & Anastasio (1968)	Misuse of analysis of covariance when treatment effect and covariate are confounded.	3
Loftin (1990)	The extreme dangers of covariance corrections.	4
Wainer (1991)	Adjusting for differential base rates: Lord's paradox again.	5
Cochran (1957)	Analysis of covariance: Its nature and uses.	6
Glass et al. (1972)	Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance.	7
Porter & Raudenbush (1987)	Analysis of covariance: Its model and use in psychological research.	8
Cox & McCullagh (1982)	Some aspects of analysis of covariance.	9
Adams et al. (1985)	Analysis of covariance as a remedy for demographic mismatch of research subject groups: Some sobering simulations.	10

Table 10: Kendall correlations and 95% credible intervals computed between the final rank based on the citation network and ranks based on mean relevance grade and number of raters.

	τ	95% credible interval
A: Rank by number of raters – Final rank	-0.032	[-0.281, 0.220]
B: Rank by relevance grade – Final rank	0.126	[-0.138, 0.363]
Rank by average of A and B – Final rank	0.077	[-0.182, 0.319]

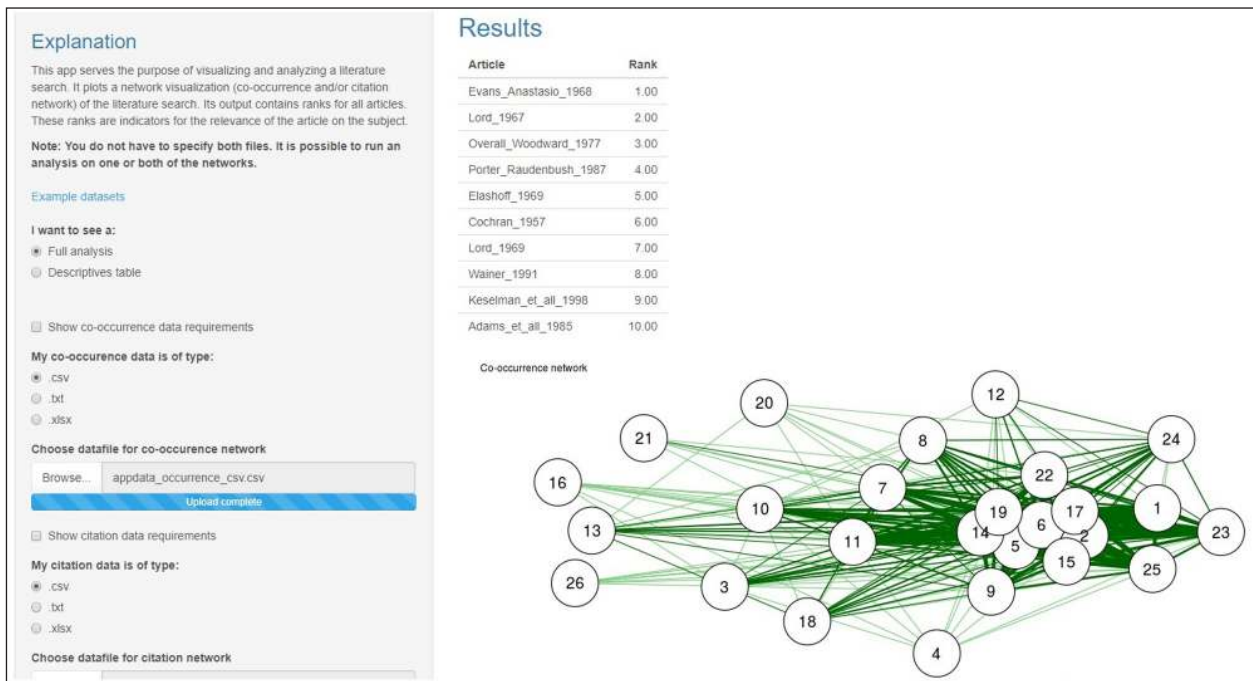


Figure 6: Screenshot of the Literature Networks web application showing a co-occurrence network with a corresponding top 10 list of most relevant articles. See online Appendix C (osf.io/bxz6c/) for a manual and walkthrough of the application.

and/or a citation network from a literature dataset and compute overall rankings using the latent data augmentation method. **Figure 6** shows a screenshot of the application. A manual containing a walkthrough and an example of the application can be found in online Appendix C (osf.io/bxz6c/).

As suggested to us in the review process, literature network models may find application in meta-analysis, where a key concern is that the contributions to the literature are based on one or two influential groups. A citation network can help visualize the interdependencies between individual contributions that would otherwise remain hidden. Specifically, the citation network can help identify different clusters of research groups, and illuminate the extent to which they consistently find the same or different results.

Finally, it should be noted that the two network models can operate sequentially instead of in parallel; specifically, citation networks could be created automatically, and the results from these networks could be used to create highly informative items for a subsequent assessment of relevance.

In sum, we have demonstrated how different network models can be applied to a dispersed scientific literature, making it easy to inspect the relationships between various articles and gauge their relative importance.

Competing Interests

The authors have no competing interests to declare.

References

Adams, K. M., Brown, G. G., & Grant, I. (1985). Analysis of covariance as a remedy for demographic mismatch of research subject groups: Some sobering

simulations. *Journal of Clinical and Experimental Neuropsychology*, 7(4), 445–462. DOI: <https://doi.org/10.1080/01688638508401276>

Albert, J. H. (1992). Bayesian estimation of the polychoric correlation coefficient. *Journal of Statistical Computation and Simulation*, 44(1–2), 47–61. DOI: <https://doi.org/10.1080/00949659208811448>

Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9(1), 91–121. DOI: <https://doi.org/10.1146/annurev-clinpsy-050212-185608>

Cochran, W. G. (1957). Analysis of covariance: Its nature and uses. *Biometrics*, 13(3), 261–281. DOI: <https://doi.org/10.2307/2527916>

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Oxford, England: Lawrence Erlbaum. DOI: <https://doi.org/10.4324/9780203774441>

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Möttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, 54, 13–29. DOI: <https://doi.org/10.1016/j.jrp.2014.07.003>

Cox, D. R., & McCullagh, P. (1982). Some aspects of analysis of covariance. *Biometrics*, 38(3), 541–561. DOI: <https://doi.org/10.2307/2530040>

Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). Qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48(4), 1–18. DOI: <https://doi.org/10.18637/jss.v048.i04>

- Elashoff, J.** (1969). Analysis of covariance: A delicate instrument. *American Educational Research Journal*, 6(3), 383–401. DOI: <https://doi.org/10.2307/1161858>
- Etz, A., Gronau, Q. F., Dablander, F., Edelsbrunner, P. A., & Baribault, B.** (2017). How to become a Bayesian in eight easy steps: An annotated reading list. *Psychonomic Bulletin & Review*. DOI: <https://doi.org/10.3758/s13423-017-1317-5>
- Evans, S. H., & Anastasio, E. J.** (1968). Misuse of analysis of covariance when treatment effect and covariate are confounded. *Psychological Bulletin*, 69(4), 225–234. DOI: <https://doi.org/10.1037/h0025666>
- Field, A.** (2013). *Discovering statistics using IBM SPSS statistics*. Sage. DOI: https://doi.org/10.1111/insr.12011_21
- Friedkin, N. E.** (1991). Theoretical foundations for centrality measures. *American Journal of Sociology*, 96(6), 1478–1504. DOI: <https://doi.org/10.1086/229694>
- Glass, G. V., Peckham, P. D., & Sanders, J. R.** (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, 42(3), 237–288. DOI: <https://doi.org/10.2307/1169991>
- Huitema, B. E.** (1980) The analysis of covariance and alternatives. Hoboken, New Jersey, John Wiley & Sons. DOI: <https://doi.org/10.1002/9781118067475>
- Jamieson, J.** (2004). Analysis of covariance (ANCOVA) with difference scores. *International Journal of Psychophysiology*, 52(3), 277–283. DOI: <https://doi.org/10.1016/j.ijpsycho.2003.12.009>
- Judd, C. M., McClelland, G. H., & Ryan, C. S.** (2011). *Data analysis: A model comparison approach*. New York: Routledge.
- Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., & Levin, J. R.** (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research*, 68(3), 350–386. DOI: <https://doi.org/10.2307/1170601>
- Loftin, L.** (1990). The extreme dangers of covariance corrections. Retrieved from: <https://eric.ed.gov/?id=ED317588>.
- Lord, F. M.** (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, 68(5), 304–305. DOI: <https://doi.org/10.1037/h0025105>
- Miller, G. A., & Chapman, J. P.** (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, 110(1), 1, 40–48. DOI: <https://doi.org/10.1037//0021-843X.110.1.40>
- Overall, J. E., & Woodward, J. A.** (1977). Common misconceptions concerning the analysis of covariance. *Multivariate Behavioral Research*, 12(2), 171–186. DOI: https://doi.org/10.1207/s15327906mbr1202_5
- Porter, A. C., & Raudenbush, S. W.** (1987). Analysis of covariance: Its model and use in psychological research. *Journal of Counseling Psychology*, 34(4), 383–392. DOI: <https://doi.org/10.1037//0022-0167.34.4.383>
- Stephenson, K., & Zelen, M.** (1989). Rethinking centrality: Methods and examples. *Social Networks*, 11(1), 1–37. DOI: [https://doi.org/10.1016/0378-8733\(89\)90016-6](https://doi.org/10.1016/0378-8733(89)90016-6)
- Storandt, M., & Hudson, W.** (1975). Misuse of analysis of covariance in aging research and some partial solutions. *Experimental Aging Research*, 1(1), 121–125. DOI: <https://doi.org/10.1080/03610737508257953>
- van Doorn, J., Ly, A., Marsman, M., & Wagenmakers, E.-J.** (2017). Bayesian estimation of Kendall's tau using a latent normal approach. *arXiv preprint arXiv:1703.01805*.
- Wainer, H.** (1991). Adjusting for differential base rates: Lord's paradox again. *Psychological Bulletin*, 109(1), 147–151. DOI: <https://doi.org/10.1037//0033-2909.109.1.147>
- White, D. R., & Borgatti, S. P.** (1994). Betweenness centrality measures for directed graphs. *Social Networks*, 16(4), 335–346. DOI: [https://doi.org/10.1016/0378-8733\(94\)90015-9](https://doi.org/10.1016/0378-8733(94)90015-9)
- Wijngaert, L., Bouwman, H., & Contractor, N.** (2014). A network approach toward literature review. *Quality & Quantity*, 48(2), 623–643. DOI: <https://doi.org/10.1007/s11135-012-9791-3>

How to cite this article: Derks, K., Burger, J., van Doorn, J., Kossakowski, J. J., Matzke, D., Atticciati, L., Beitner, J., Benzesin, V., de Bruijn, A. L., Cohen, T. R. H., Cordesius, E. P. A., van Dekken, M., Delvendahl, N., Dobbelaar, S., Groenendijk, E. R., Hermans, M. E., Hiekkaranta, A. P., Hoekstra, R. H. A., Hoffmann, A. M., Hogenboom, S. A. M., Kahveci, S., Karaban, I. J., Kevenaer, S. T., te Koppele, J. L., Kramer, A.-W., Kroon, E., Kucharský, Š., Lieuw-On, R., Lunansky, G., Matzen, T. P., Meijer, A., Nieper, A., de Nooij, L., Poelstra, L., van der Putten, W. J., Sarafoglou, A., Schaaf, J. V., van de Schraaf, S. A. J., van Schuppen, S., Schutte, M. H. M., Seibold, M., Slagter, S. K., Snoek, A. C., Stracke, S., Tamimy, Z., Timmers, B., Tran, H., Uduwa-Vidanalage, E. S., Vergeer, L., Vossoughi, L., Yücel, D. E., and Wagenmakers, E.-J. (2018). Network Models to Organize a Dispersed Literature: The Case of Misunderstanding Analysis of Covariance. *Journal of European Psychology Students*, 9(1), 48–57, DOI: <https://doi.org/10.5334/jeps.458>

Published: 12 December 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 Unported License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Journal of European Psychology Students is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS