

# Supplementary Information

## for

### Mapping the field of cultural evolutionary theory and methods in archaeology using bibliometric methods

David N. Matzig<sup>1</sup>✉ ORCID: 0000-0001-7349-5401

Clemens Schmid<sup>2,3</sup> ORCID: 0000-0003-3448-5715

Felix Riede<sup>1</sup> ORCID: 0000-0002-4879-7157

<sup>1</sup> Department of Archaeology and Heritage Studies, Aarhus University, Moesgård Allé 20, 8270 Højbjerg, Denmark

<sup>2</sup> Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig 04103, Germany

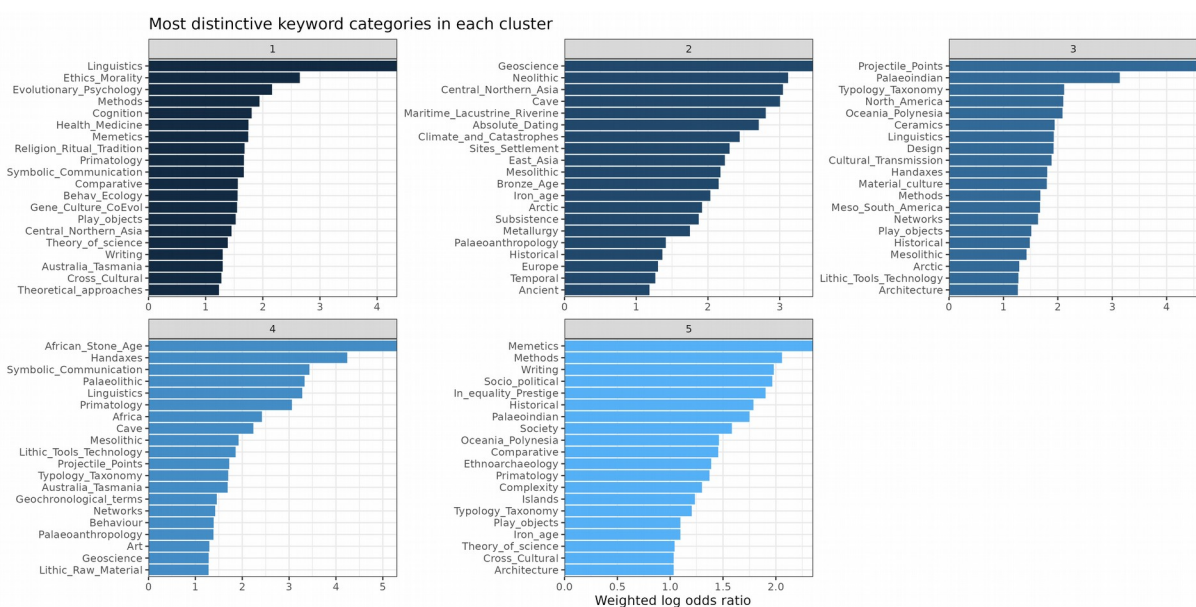
<sup>3</sup> International Max Planck Research School for the Science of Human History, Max Planck Institute for Geoanthropology (formerly known as Max Planck Institute for the Science of Human History), Kahlaische Str. 10, Jena 07745, Germany

✉ Correspondence: David N. Matzig <david.matzig@cas.au.dk>

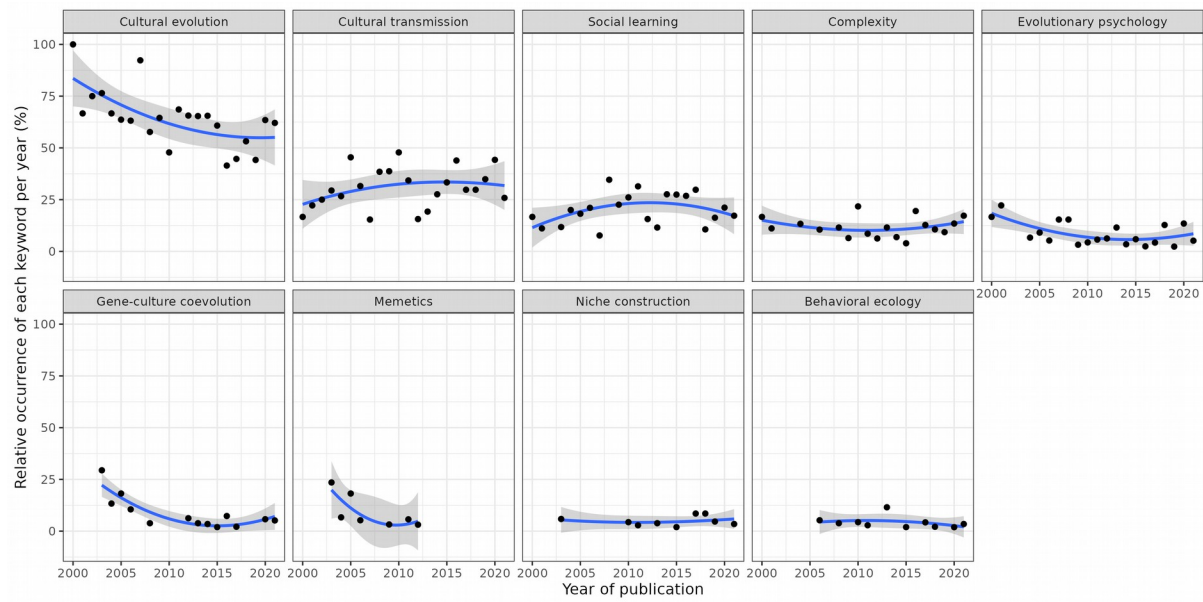
The exact search query for the WoS data base was

*(TS=(archaeolog\*) AND ALL=("cultur\* evolutio\*)) OR (TS=(archaeolog\*) AND TS=("cultur\* transmissio\*)) OR (TS=(anthropology) AND ALL=("cultur\* evolutio\*)) OR (TS=(anthropology) AND ALL=("cultur\* transmissio\*))*

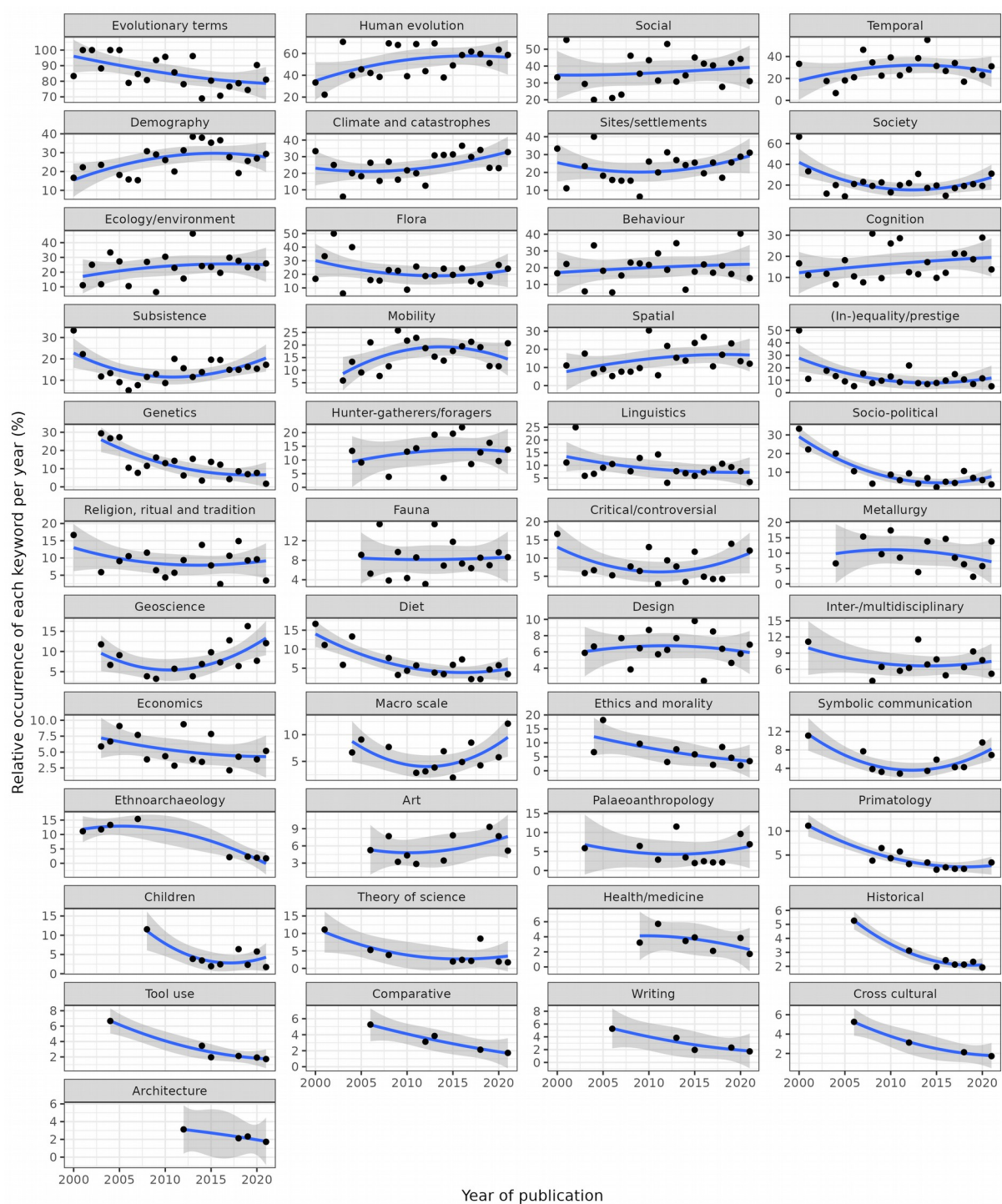
for the time frame 1900-01-01 to 2021-12-31, which, as of March 2022, resulted in 677 raw results. Only 674 of them actually are from this time frame. This syntax makes sure to capture the search terms in all possible fields where they potentially appear with different notations (i.e. "cultur\* evolutio\*" detects “culture evolution” as well as “cultural evolutionary”) for the WoS topic categories of *archaeology* and *anthropology*.



Supplementary Figure S1: *Weighted log odds ratio for clusters 1-5. The keywords (y-axis) are sorted according to their weighted log odds ratio (x-axis), where high values represent strong cluster association.*



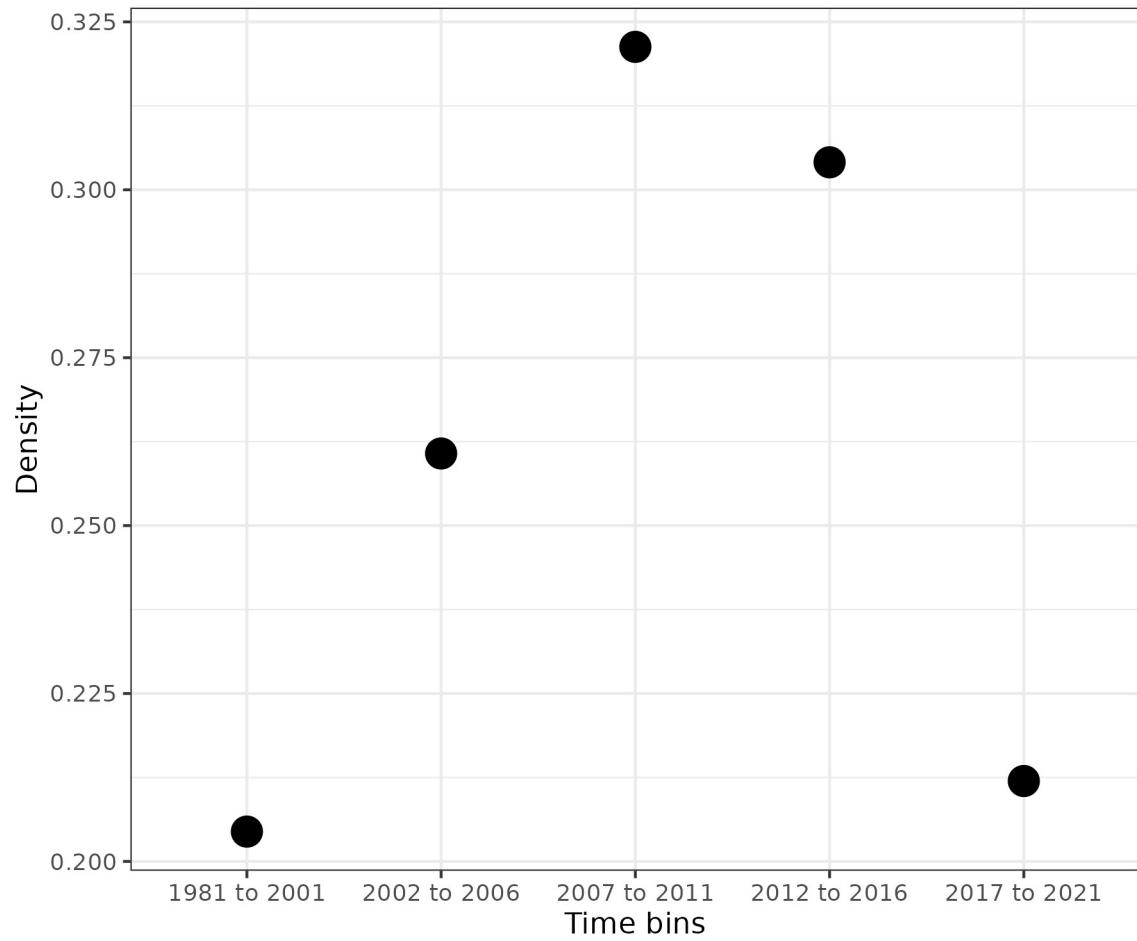
Supplementary Figure S2: *Relative appearance of keywords related to the thesaurified Cultural Evolutionary Theory (CET) meta-category treated in the whole corpus per year.*



Supplementary Figure S3: *Relative appearance of keywords related to thesaurified topics treated in the whole corpus per year. The y-axis is scaled dynamically for each subplot. A comparison of trends can only be done within categories but not between categories.*

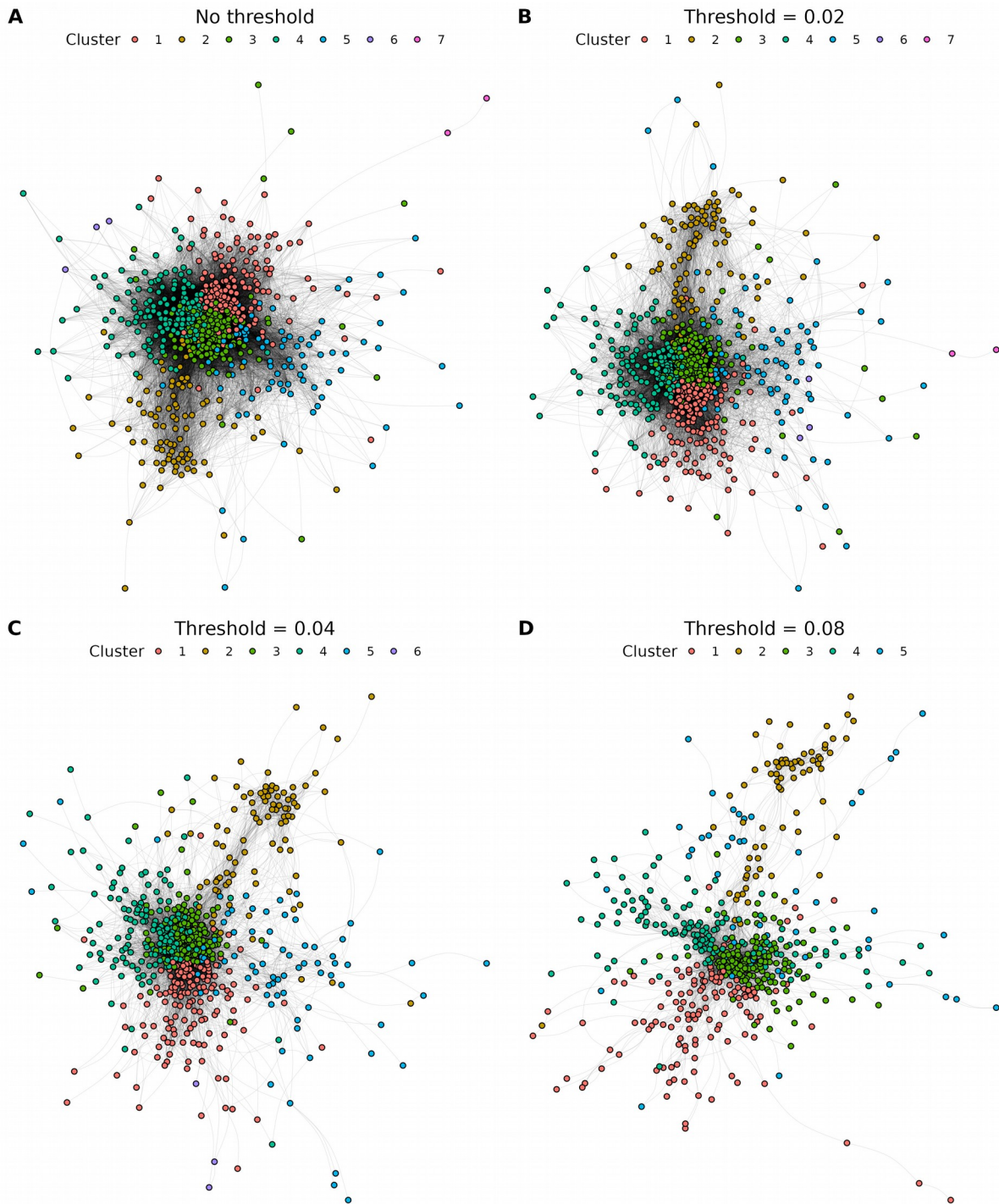


Supplementary Figure S4: *Bibliographic coupling network split into each of the five separate time bins. The clusters are coded in the same individual colours as Fig. 5 in the main text and were retrieved from the complete data set ( $N=629$ ) without time binning. The graph was visualised using the Fruchterman-Reingold algorithm in the ggraph package (Pedersen 2021).*



Supplementary Figure S5: *Bibliographic coupling network density (ratio of existing edges to all possible edges) for each of the separate time bins. Here, network density is used as a proxy for the conformity/diversity within the network.*





Supplementary Figure S6: *Bibliographic coupling network with different thresholds applied to the underlying cosine similarity matrix. A is the same graph as Figure 5 in the main article with no threshold applied, B, C, and D have a threshold of 0.02, 0.04, and 0.08 respectively, meaning that edges below said values were deleted. The clusters retrieved by the Louvain algorithm are differentiated by colour. For B, C, and D, the article-cluster associations were taken from network A. The graph was visualised using the Fruchterman-Reingold method in the ggraph package (Pedersen 2021).*

Supplementary Table S1: *The result (p-values) of the pairwise comparisons of the closeness centrality between clusters using a t-tests with pooled standard deviation and Bonferroni adjustment method.*

<b>Cluster</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>2</b>	0.29518	-	-	-	-	-
<b>3</b>	1.0e-04	3.3e-09	-	-	-	-
<b>4</b>	3.3e-06	6.4e-11	1	-	-	-
<b>5</b>	1	0.02024	0.20939	0.03595	-	-
<b>6</b>	1	1	0.14502	0.09743	0.79305	-
<b>7</b>	0.00016	0.00124	3.9e-06	2.2e-06	5.5e-05	0.37105