

Quantifying the dynamics of nearly 100 years of dominance hierarchy research

Elizabeth A. Hobson^{1,*}

¹*Department of Biological Sciences, University of Cincinnati, Cincinnati, OH USA*

^{*}*Corresponding author; e-mail: elizabeth.hobson@uc.edu*

1 **Abstract**

2
3 Dominance hierarchies have been studied for almost 100 years. The science of science approach
4 used here provides high-level insight into how the dynamics of dominance hierarchy research have
5 shifted over this long timescale. To summarize these patterns, I extracted publication metadata
6 using a Google Scholar search for the phrase “dominance hierarchy”, resulting in over 26,000
7 publications. I used text mining approaches to assess patterns in three areas: (1) general pat-
8 terns in publication frequency and rate, (2) dynamics of term usage, and (3) term co-occurrence
9 in publications across the history of the field. While the overall number of publications per decade
10 continues to rise, the percent growth rate has fallen in recent years, demonstrating that although
11 there is sustained interest in dominance hierarchies, the field is no longer experiencing the ex-
12 plosive growth it showed in earlier decades. Results from title term co-occurrence networks and
13 community structure show that the different subfields of dominance hierarchy research were most
14 strongly separated early in the field’s history while modern research shows more evidence for co-
15hesion and a lack of distinct term community boundaries. These methods provide a general view
16 of the history of research on dominance hierarchies and can be applied to other fields or search
17 terms to gain broad synthetic insight into patterns of interest, especially in fields with large bodies
18 of literature.

19

20 **Keywords**

21 Science of science, dominance hierarchy, text mining

22 **Introduction**

23 Competition is nearly ubiquitous in situations where resources are limited and contested. Be-
24 cause of this, conflict is inevitable in most social groups, leading to increased access to these

25 resources for some individuals and decreased access for others. In many social species, this
26 competition leads to the emergence of group dominance hierarchies, which can help make so-
27 cial life more structured and predictable and regulate overall conflict. In nearly 100 years of
28 research on dominance, scientists have documented the presence of hierarchies that structure
29 social conflict in a wide range of species [24; 23]: groups of ants, fish, lizards, geese, parrots, ele-
30 phants, hyenas, primates, and many species in between form groups with detectable hierarchies,
31 where individuals within the groups can be assigned ranks. Much research has also established
32 that these hierarchies also matter to individuals: higher-ranked individuals often benefit from im-
33 proved health or access to resources, more reproductive opportunities, more offspring, or greater
34 longevity [20; 2; 17; 21].

35 Scientific progress itself is a social process, with new research continually building on the founda-
36 tions of previous work. A science of science approach can be used to synthesize the history of
37 a large and active field like dominance research. This approach complements more typical liter-
38 ature reviews with a “big data” perspective on publishing patterns and topics in the field. There
39 have been too many papers published on dominance hierarchies in too many subfields for it to be
40 feasible to read and synthesize the entire body of work. A traditional literature review may also
41 be unintentionally biased towards certain subtopics, study species, or subfields. The quantitative
42 approach used here allowed me to summarize a much bigger body of literature to identify general
43 patterns and provide an overall summary of how areas of focus and study have changed over the
44 course of nearly 100 years of research since the original publication describing “peck order” in
45 chickens [22]. However, it is important to note that what we gain from this high-level perspective
46 is balanced by the absence of important syntheses of knowledge that are part of good literature
47 reviews (see other papers in this special issue for this perspective).

48 Here, my goal was to quantify general patterns of interest in this long-running field of research.
49 Specifically, I focused on three main aspects: (1) general patterns in overall publication frequency
50 and rate, (2) dynamics of changes in term usage in titles, and (3) how terms co-occur in publica-
51 tion titles across the history of the field. I used a combination of data scraping, text mining, and network
52 analysis to quantify these patterns. This approach provides a broad review of the history of study
53 of dominance hierarchies to better understand where and how researchers have focused their
54 scientific efforts.

55 **Methods**

56 ***Data collection and processing***

57 I used the program “Publish or Perish” [27; 11] to scrape Google Scholar for publications using the
58 search phrase [“dominance hierarchy”]. The Google Scholar algorithm returns publications which
59 match the search phrase anywhere in the searched documents (author, title, source, abstract,
60 references, etc.) [11]. While the full Google Scholar search algorithm is not publicly available, the
61 algorithm generally works by considering the full text of each document, the publication venue, the
62 authors, and the recency and frequency with which other papers have cited it to rank all publica-
63 tions and return the top 1000 results. I used Google Scholar as the search engine because it has
64 wider coverage and returns more publications than many other searches (i.e. Web of Science,
65 [12]). It also returns a broader array of publication types including “grey literature” like theses, con-
66 ference proceedings, white papers, and preprints [10], the inclusion of which is one way to reduce

67 biases in literature results [9]. While using Google Scholar is beneficial for its wider reach, there
68 are some important limitations: it does not have strong quality control processes so data are gen-
69 erally not as clean as the more limited output from other sources [12], resulting in some records
70 with missing data, corrupted text, erroneous publication years, etc. Results, including titles, are
71 also sometimes truncated if they are too long [11].

72 The use of “dominance hierarchy” as the search phrase also has limitations. First, searching for
73 this term heavily influences search results towards English-language literature, which can limit
74 and bias large literature reviews [19; 16]. While searching for a translation of dominance hierarchy
75 would have been feasible, matching up all the terms in titles for the analyses here would have
76 required a large amount of translation and unfortunately was not feasible for this project. The
77 second limitation of this search phrase is that it may miss conceptually similar papers, for example,
78 papers focusing on rank, agonistic contests, or contest outcomes, which may not include the
79 phrase “dominance hierarchy” within the paper. This issue is likely reduced by the choice of the
80 Google Scholar search algorithm, which considers the full text of papers, rather than just the terms
81 in the titles or keywords, but is an important consideration.

82 Using Publish or Perish, I collected the titles of the top 1000 Google Scholar results of publications
83 returned with the search phrase for each year (searches conducted from 2021-04-14 to 2021-
84 04-17). Any records missing a title or a valid publication year were excluded from all further
85 analyses. I also excluded any title with a publication year prior to 1920 to pin the start of analyses
86 to the decade in which the foundational paper for the field, by Schjelderup-Ebbe, was published in
87 1922 [22]. Analyses were pooled by decade to facilitate identification of general patterns.

88 ***Quantifying publication trends over time***

89 To quantify how publications trends involving the search phrase “dominance hierarchy” have changed
90 over time, I counted the number of publications in the database for each decade. I then quantified
91 the lagged change in publication number by comparing each decade’s total publications to the
92 previous decade. Finally, I quantified the decade over decade percent growth in publications by
93 dividing the lagged change in publication number by the number of publications in the previous
94 decade and converting to a percent.

95 ***Estimating dynamics of term usage in publication titles***

96 All titles in the database were cleaned prior to text analyses using the R packages tm and stringr [28;
97 6]: punctuation, numbers, and whitespaces were removed, and all was text converted to lower-
98 case. I removed all stopwords (extremely common words like a, is, the, etc.) using the snowball
99 stopword list in the R tidytext package [25]. Prior to full analyses, all terms in the dataset were also
100 stemmed using the the tm package. Stemming reduces instances of similar words (e.g., “ecolog-
101 ical” and “ecology” both reduce to the stemmed term “ecolog”). Limited further stemming efforts
102 for commonly-used words were conducted by hand (e.g., both “behavior” and “behaviour” were
103 retained as separate terms following stemming in the tm package; I collapsed these to “behav-
104 ior”). These cleaned and stemmed terms formed the initial corpus and are referred to here as “title
105 terms”.

106 To estimate title term use dynamics, I filtered the terms in the title corpus to retain only the title
107 terms used in at least 25 publications in the database. This process excluded terms that were

108 rarely used but also helped exclude any terms that were reported by Google Scholar in a some-
109 what corrupted format (for example, punctuation was sometimes introduced erroneously within
110 words in the search results, some words were split by truncation, etc.). This dataset of non-rare
111 title terms was used in all further analyses. I then determined whether each title term was present
112 or absent in each document's title (this corrects for cases where a particular title term was used
113 multiple times within a single title). I used presence data to then find the total number of publica-
114 tions per decade where each title term was present in the title.

115 To determine how title terms were used from decade to decade, I quantified the Shannon diversity
116 of title term presence in publications in each decade, using the frequency with which title terms
117 were present in titles compared to the total number of unique title terms. I also quantified the
118 number of novel title terms per decade, where "novelty" was defined as a unique title term which
119 appeared in a title where it was not present in earlier decades in the corpus. Finally, using the
120 entire history of dominance hierarchy title term corpus, I found the percent of total title terms that
121 were present in titles in each decade.

122 To measure decade to decade similarity in title term usage, I compared the presence of title terms
123 in each decade to the presence of title terms in all other decades. I then found the number of title
124 terms that were present in both decades, the percent of words present in both decades, and the
125 percent similarity of each decade compared to the decade it was most similar to. This analysis
126 helped determine "hot spots" of title term use similarity in documents published across different
127 decades, where the more similar term use was, the more similar those decades would be when
128 compared.

129 ***Determining differences in title term co-occurrence***

130 Knowing that a certain single term was present in titles in a particular decade is helpful for as-
131 sessing when terms emerge and are popular. However, even more insight into how the focus of
132 dominance hierarchy research has changed can come from quantifying how pairs of terms co-
133 occur in documents through the history of the field. To do this, I built networks based on title term
134 co-occurrences in publication titles for each decade. Title terms formed the nodes in these net-
135 works and were connected by edges if the two terms were used in the same title within the same
136 decade. The strength of these edges showed how often terms co-occurred in titles at the decade
137 scale.

138 From these title term co-occurrence networks, I could then determine if particular title terms in
139 these networks could be assigned to different network communities, depending on how each term
140 was connected to others. Inspired by methods used to reconstruct the cultural evolution of a
141 music genre [29], I used the R package igraph and fastgreedy community detection [3; 4], to
142 identify communities of co-occurring title terms by decade, where title terms often found together
143 in publication titles in the same decade were more likely to be assigned to the same title term co-
144 occurrence community. This approach differs from common bibliometric network analyses, which
145 often focus on communities driven by shared co-authorship patterns; here, the focus is solely
146 on the content of titles rather than the identity of the authors. I calculated the modularity of title
147 term co-occurrence networks in each decade, which measures how well the community-detection
148 algorithm partitions a network into communities [3]. Finally, for each decade, I quantified how title
149 term co-occurrence communities were interconnected or separated from each other [18]. To do

150 this, I calculated a cohesion index [29] to represent the ratio of within-community edges compared
151 to connections between title terms assigned to different title term co-occurrence communities,
152 using unweighted binary edges. A cohesion index near 0 indicates that few connections between
153 terms assigned to different communities exist; a value near 1 indicates that most connections
154 occur between terms assigned to different communities.

155 Finally, because community detection was done independently by decade, I needed a way to track
156 similarity in use of title terms across decades. Title term co-occurrence communities detected in
157 each decade were assigned numerical codes, but “Community 1” in one decade is not necessarily
158 comprised of a similar assortment of title terms in the next decade - in other words, there was no
159 consistent naming continuity in communities across decades. To determine cross-decade commu-
160 nity similarity, I identified which title term co-occurrence communities in one decade were the most
161 similar in term composition to a community detected in the next decade. I used Jaccard similarity,
162 which measures overlap in membership between sets as the size of the intersection divided by
163 the size of the union of two sets. In my case, I used Jaccard similarity to find the proportion of
164 title terms found in a title term co-occurrence community in one decade compared to the title term
165 compositions of title term co-occurrence communities in the next decade. The similarity measure
166 ranges from 0 (no overlap in terms) to 1 (exactly the same terms present in both communities).
167 This analysis allowed me to look at how consistently terms were assigned together as a group in
168 one decade compared to the next decade.

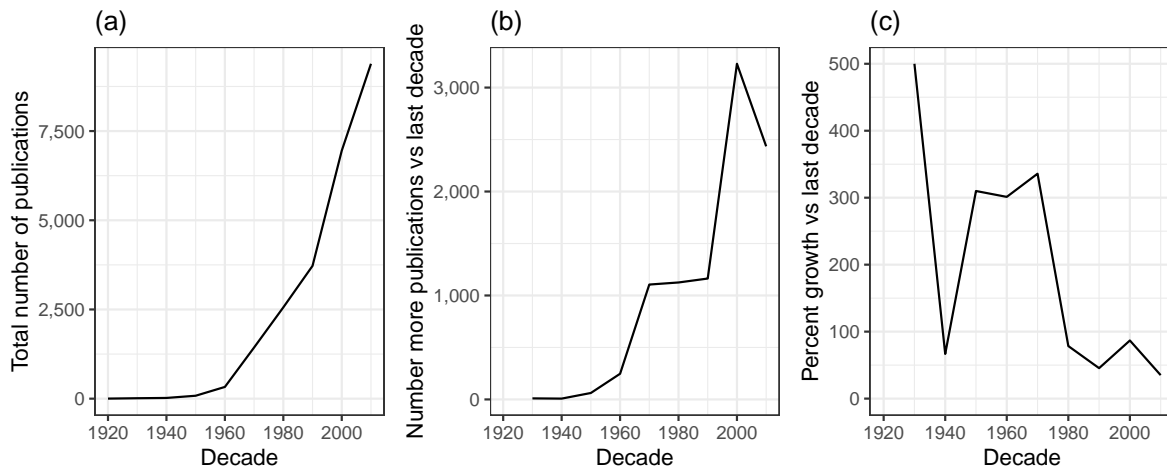
169 **Results and Discussion**

170 ***Quantifying publication trends over time***

171 Analyzing publication trends can provide insight into overall interest in a field of study. Research
172 on dominance hierarchies has resulted in an impressive number of publications in nearly 100
173 years of research, with over 26,000 publications in the scraped dataset which were returned from
174 a keyword search for “dominance hierarchy”. The Google Scholar algorithm returns a maximum of
175 1000 results per query, so queries were split by time periods as small as a single year to maximize
176 results (2011 was the only year where 1000 results were returned; all other years were below this
177 thresholding limit and represent complete search results). After data cleaning (which excluded
178 any publication missing a publication year and/or a title, or with a publication date prior to 1922),
179 25,219 publications were retained for the analyses.

180 Figure 1 shows three views of publication trends by decade. While publication numbers have risen
181 from decade to decade during the entire history of the field (Fig. 1a), it is also important to account
182 for the overall increase in modern publication rates. Quantifying how the number of publications in
183 one decade compares to the number of publications in the last decade helps somewhat normal-
184 ize for this general increase in overall publications in modern science and helps better visualize
185 changes in interest on a decade-by-decade scale. For example, the greatest increase in numbers
186 of papers in a decade compared to the last decade occurred during the 2000’s. 2010 was the first
187 decade in which this explosive growth rate decreased: fewer additional papers were published in
188 the 2010’s compared to the increased number of publications when comparing the 2000’s to the
189 1990’s. The decade-over-decade percent growth rate provides different insight into publication
190 trends, with the highest percent growth seen in the 1930’s compared to the 1920’s. This high
191 growth rate indicates just how quickly the initial number of publications grew when compared to

Figure 1: Publications by decade showing (a) total publications for each decade, (b) the number of publications in each decade compared to the previous decade, and (c) the percent growth in number of publications compared to the previous decade.



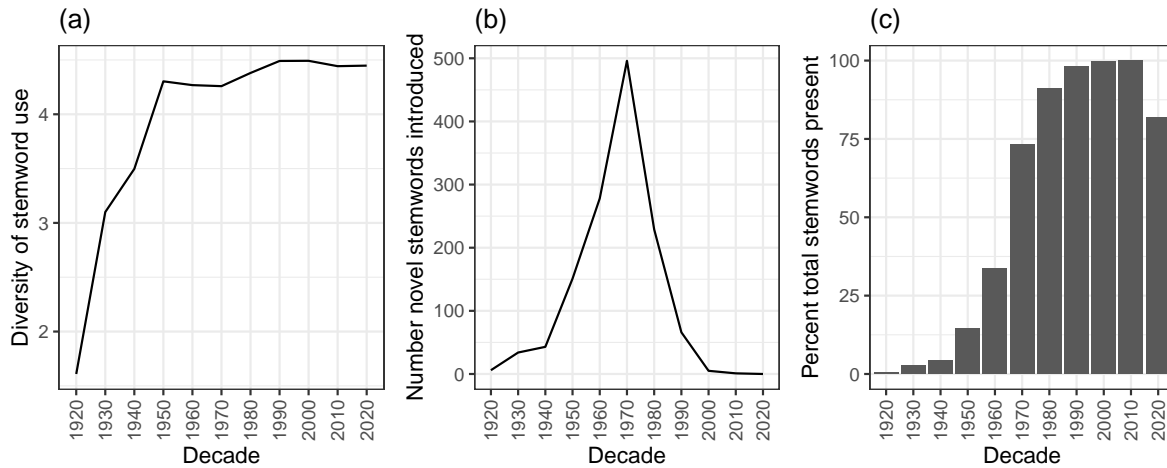
192 the “founding” of dominance hierarchy research in the 1920’s. We see another peak in percent
 193 growth rates in the 1950’s to the 1970’s with decade over decade growth rates around 300%.
 194 These publication trends provide strong evidence for sustained interest in dominance hierarchy
 195 research despite nearly 100 years of study, but do indicate that in the most recent decades, the
 196 earlier explosive growth in numbers of publications has tapered off.

197 *Estimating dynamics of term usage*

198 Of 22,406 unique title terms identified via stemming words used in titles in the entire database,
 199 only 1295 were used in at least 25 separate publications. I used these more commonly-used terms
 200 to estimate dynamics of term usage. This 25-publication threshold filtered out a large proportion
 201 of potential title terms, but allowed me to focus on ones that were more-used and was an easy
 202 way to exclude the many corrupted terms found in the Google Scholar search.

203 While general patterns in the number of publications per decade provide evidence of sustained
 204 interest in dominance hierarchy research, how terms are used each decade provides insight into
 205 areas of focus for research efforts. I found that the overall diversity of title term use per decade
 206 increased sharply up to 1950, then continued to increase at a slower rate until 1990 (Fig. 2a).
 207 The diversity of title term use has been relatively stable from 1990 to 2020 and reflects the highest
 208 diversity period in the history of dominance hierarchy research. This high diversity period coincides
 209 with high overall numbers of publications during these decades, as seen in Figure 1. Interestingly,
 210 despite very few years of publications so far in the 2020’s, publications from January 2020-April
 211 2021 already share the high diversity of title term use seen across the whole 2010 decade. Overall
 212 patterns of term diversity and network measures (see below) are unlikely to result from decade-to-
 213 decade changes in publications norms such as title lengths; title length was relatively consistent
 214 across the dataset with the median number of title terms per title per decade ranging from 3 to 6
 215 and an average median number of title terms per publication per decade of 4.77.

Figure 2: Title term usage in titles by decade showing (a) Shannon diversity in title term use in each decade (with diversity calculated on total number of publications using each title term per decade), (b) the number of novel title term introduced in each decade, and (c) the percent of title terms in the entire corpus that were used in titles in each decade.



216 The number of novel title terms introduced in publication titles peaked sharply in 1970 (Fig. 2b).
 217 On a decade by decade basis, Figure 2c shows how terms present across the entire history of
 218 dominance hierarchy research are used in publication titles: by 1970, nearly 75% of all title terms
 219 in the corpus were in use.

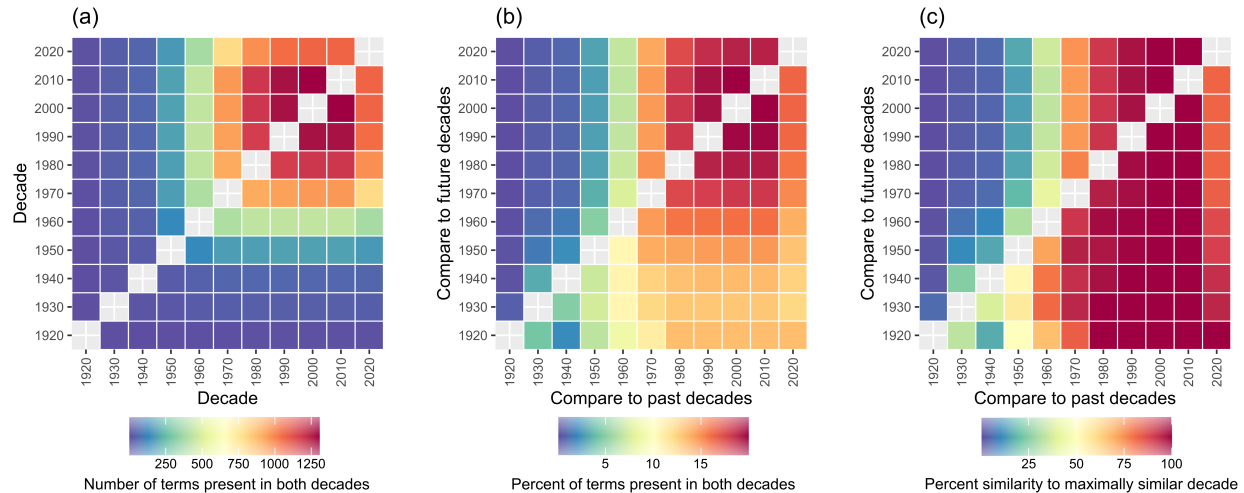
220 Each decade can also be compared to other decades in the dataset to determine the levels of
 221 similarity in title term use in titles over time (Fig. 3). The number of title terms present in both
 222 decades peaked when comparing 1990, 2000, and 2010. Breaking these patterns down by the
 223 percent of terms present in both decades (compared to the number of terms present in either
 224 decade) shows an even larger hotspot of similarity that highlights how similar term use in 1980-
 225 2020 has been to other decades within that same time span. Comparing term use in each decade
 226 to the decade with which term use is most similar shows an even wider hotspot coinciding with all
 227 decades compared to 1990-2010; this hotspot also coincides with data in Figure 2c, showing that
 228 nearly 100% of title terms used in the entire historical corpus were used in 1990-2010.

229 **Determining patterns in title term co-occurrences**

230 In addition to quantifying overall publication trends and the use of single terms in titles, the co-
 231 occurrence of terms in titles can provide insight into areas of focus for dominance hierarchy re-
 232 search and how the use of pairs of terms in titles has changed over time. Using network measures
 233 (eigenvector centrality and community detection) I found each term's community membership in
 234 each decade, with community membership determined by how each term co-occurred in titles
 235 with other terms. These term co-occurrences and community assignments across decades are
 236 depicted as wordclouds in Figure 4, visualized with the R package 'wordcloud' [7].

237 Figure 5 shows how terms can shift in how they co-occur in titles over time. In the figure, title term
 238 use and title term co-occurrence community membership are shown for the top three most-used

Figure 3: Title term usage in publications showing “hotspots” of similarity across decades: (a) similarity in the raw number of title terms present in both decades, (b) scaled similarity showing the percent of title terms present in decades on the y-axis compared the title terms present in decades on the x-axis, and (c) relative similarity showing each decade’s similarity scaled by maximum similarity to decades on the y-axis, with maximum similarity shown in red.



239 terms (“behavior”, “social”, and “domin”). In 1930, all three title terms were assigned to separate
 240 communities, while in 1940, the title terms “behavior” and “social” were grouped together in
 241 Community 1 and “domin” was assigned to Community 2. In 1950 this pattern changed again as
 242 “behavior” was assigned to Community 1 and “social” and “domin” were both assigned to Commu-
 243 nity 2. This pattern illustrates that title term co-occurrence community membership composition
 244 changed from year to year, with even the top most-used terms changing in how they were grouped.
 245 Interestingly, “social” and “behavior” both appeared in a higher proportion of titles in almost every
 246 decade than “domin”, which could be an indication of a long-standing trend in dominance hierarchy
 247 research for placing dominance within a social and behavioral context.

248 While the composition of title term co-occurrence communities has changed over time, the dis-
 249 tinctness of borders between these communities has also changed. Figure 6 shows how title term
 250 co-occurrence networks were much more modular early in dominance hierarchy research, but
 251 have decreased in modularity in recent years, indicating that terms have become less-strictly co-
 252 occurring with specific other terms in titles. The number of detected communities has also shifted
 253 from a small number of communities to a peak of 12 detected communities in 1950, followed by
 254 a gradual decline to fewer communities in modern dominance hierarchy research (it is uncertain
 255 if the rise in community number in the 2020’s will persist as more papers are published in this
 256 decade so those results should be treated with caution).

257 Within decades, the percent of edges between title terms assigned to the same community com-
 258 pared to title terms assigned to different communities has changed over time. Connections be-
 259 tween terms assigned to different communities have increased over the history of dominance
 260 hierarchy research to plateau in modern times at about 50%, indicating a balance between within-
 261 community and outside-community edges. This connectivity pattern contributed to the decrease

Figure 4: Wordclouds showing title term community assignment across decades. Text size indicates each term’s eigenvector centrality in the title term co-occurrence network for that decade, with higher centrality terms printed in larger text. Text color indicates each term’s assignment to a community in each decade; terms in the same color within the same decade were assigned to the same title term co-occurrence community. To improve legibility, wordclouds show (at most) the top 10 most central terms per community per decade.

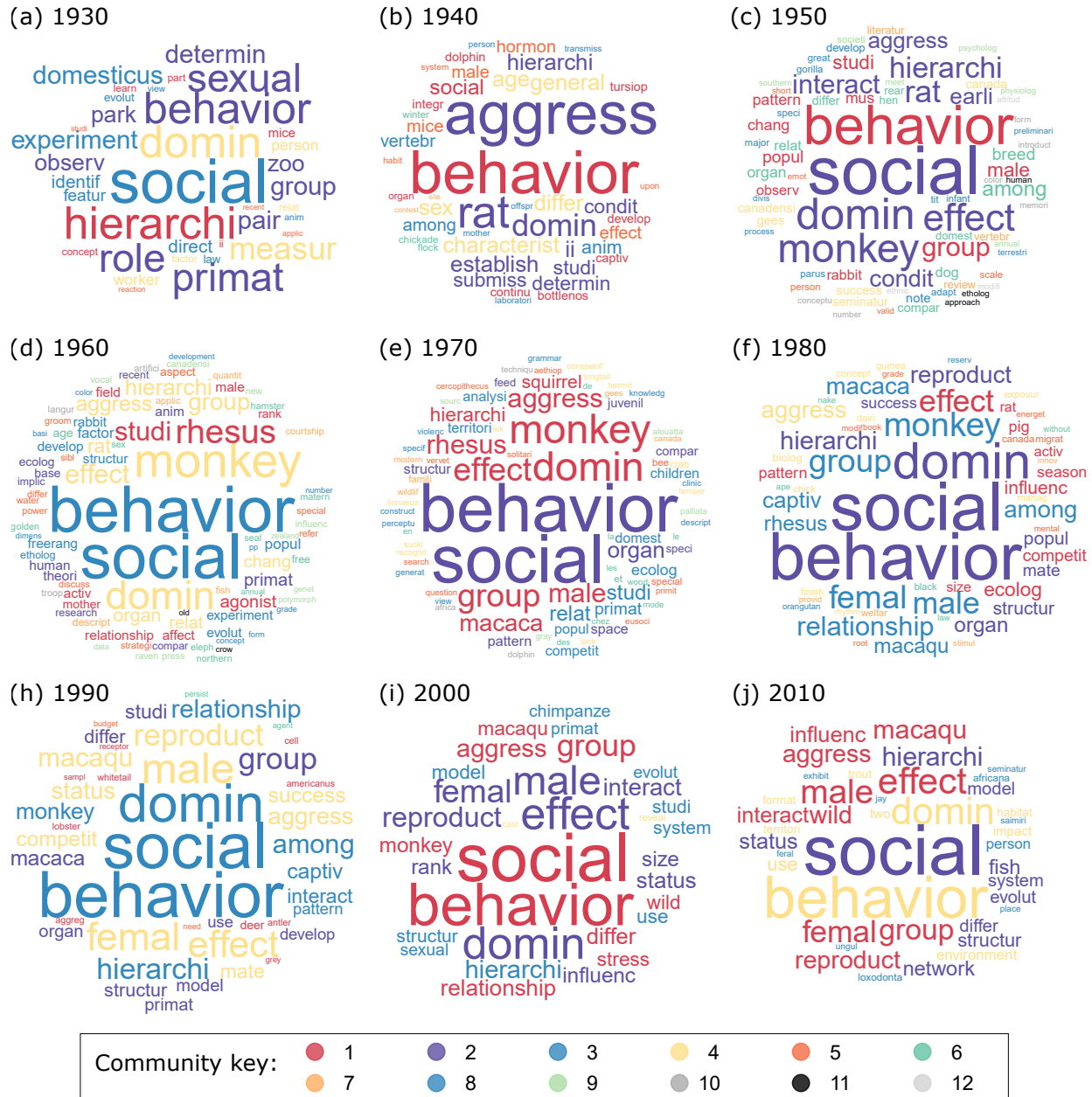
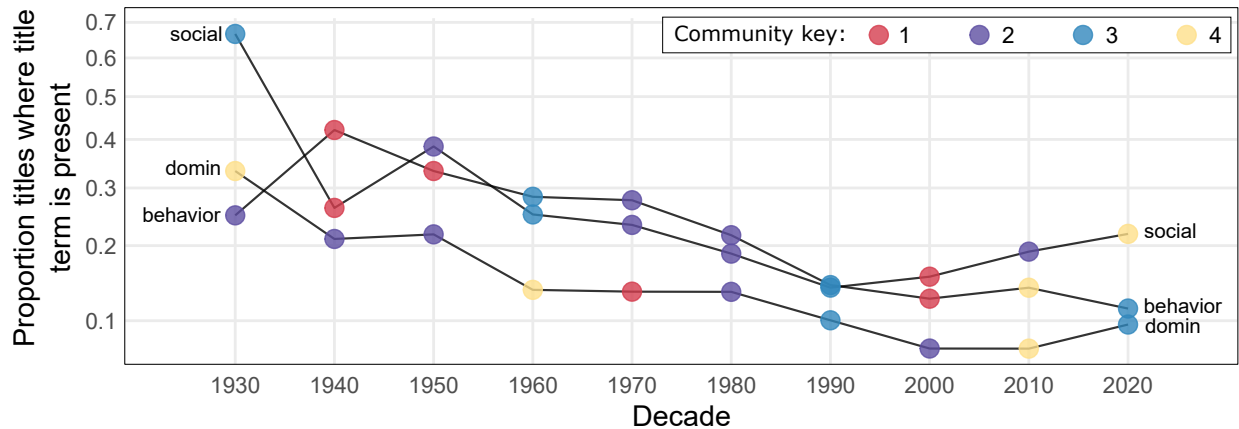


Figure 5: Title term use and co-occurrence community membership over time for the top three most-used title terms: “behavior”, “social”, and “domin”. Points show the proportion of titles containing each title term per decade; point color indicates title term co-occurrence community membership.



262 in overall modularity as communities became more interconnected as well as the reduction in
 263 the total number of detectable communities. Whether this decrease in modularity comes from
 264 more integrative studies or potentially from a increase in cross-disciplinary work that makes more
 265 connections to research across subfields remains to be seen. Future work focused on citation
 266 networks, analyzing how authors cite each other’s work, could provide important insight into this
 267 process in the field of dominance hierarchy research (e.g., [8]).

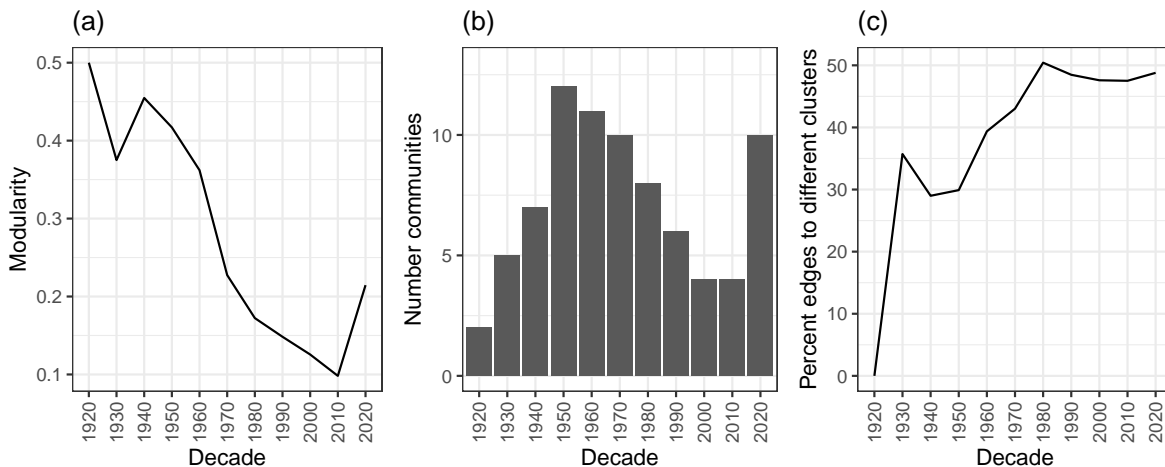
268 Across decades, title term community continuity was highest from 1990 to 2010, with an average
 269 maximal Jaccard similarity of over 25% shared title terms for a community in one decade compared
 270 to the set of title terms in the most-similar community in the next decade. This period of more
 271 stable title term assignments to similar communities co-coincided with a period of fewer overall
 272 identified communities. This result also shows that even in decades with few overall communities,
 273 the majority of title terms in any one community were not assigned together in the same community
 274 in the next decade, indicating a high level of remixing of title term use in publication titles.

275 Conclusions

276 The science of science perspective used here provided insight into general publication trends, title
 277 term use, and term co-occurrence in titles returned from a search for “dominance hierarchy” in
 278 publications across the nearly 100 years of dominance hierarchy research. From this, we can infer
 279 how investment in publishing dominance hierarchy research has changed, but also how connect-
 280 ivity between different subfields and topics has shifted.

281 An important limitation of this approach is that the analysis only considers terms used in titles of
 282 publications. This analysis obviously cannot capture the complexity of how topics are treated in the
 283 full text of these publications, so cannot provide a detailed account of exactly how research trends
 284 or concepts have shifted over the history of dominance hierarchy research. However, this sum-
 285 mary demonstrates that dominance research has had sustained interest over its long history and

Figure 6: Title term co-occurrence network summaries by decade: (a) modularity over time, (b) number of communities detected in each decade, and (c) percent of edges connecting a title term node in one community to a title term node in a different community.



286 the evidence, especially the high connectivity between what could have been isolated subfields,
 287 provides a high-level perspective on historical and modern trends in the science of hierarchies.

288 An open question among researchers working on dominance hierarchies is whether we have
 289 “solved” dominance. Informally, and depending on who you talk to at conferences, the question
 290 of dominance has been “solved” in the 1960’s, the 1970’s, the 1980’s, or the 2000’s. However,
 291 the sustained interest and investment in new publications demonstrated here, as well as greater
 292 cohesion and cross-community connections, suggests that there is still much interest in all that we
 293 still have to learn about dominance hierarchies.

294 Recent dominance studies may be moving to a new stage of research focus, particularly if we have
 295 solved some of the more basic hierarchy questions. In particular, new genetic methods (e.g., [1],
 296 this issue), computational approaches [14], and a focus on the information contained in both net-
 297 works of aggression and rank within social groups [14; 13] provide many new avenues for novel
 298 insight into animal sociality. Both theoretical work (e.g., [15; 5] and empirical work [14] have also
 299 recently suggested that rank acquisition can be remarkably sensitive to stochastic events. This
 300 new work has the potential to enrich our understanding of how rank forms and is maintained in
 301 different groups (see Tibbetts, this issue). Other empirical work has shown that “rule-breaking”
 302 via coalition formation can cause disruptions to expected rank inheritance patterns [26], and that
 303 these dynastic changes can gain momentum and persist despite the lack of underlying character-
 304 istics or quality to differentiate these individuals from less-successful ones in the group. Finally,
 305 comparative analyses across species have strong potential to advance our understanding of dom-
 306 inance hierarchy structures. New compilations of dominance data across species, such as in the
 307 R package DomArchive (see Strauss et al., this issue), provide easier access to historical datasets
 308 which researchers can use to test hypotheses about dominance and how the social and cognitive
 309 features required for dominance to emerge may have evolved. These datasets can also form the
 310 basis for new analyses to understand how rank influences individual health and how competition
 311 can influence the outcome of fitness-related traits more broadly.

312 The integration of new tools as well as new, more complex ways of studying the decisions animals
313 make about who, when, and how they fight each other, and the consequences of different conflict
314 management styles, provide a strong foundation for the next 100 years of dominance hierarchy
315 research.

References

- [1] Anderson, J. A., Lea, A. J., Voyles, T. N., Akinyi, M. Y., Nyakundi, R., Ochola, L., Omondi, M., Nyundo, F., Zhang, Y. *et al.* 2021 Distinct gene regulatory signatures of dominance rank and social bond strength in wild baboons. *bioRxiv*, p. 2021.05.31.446340. (doi:10.1101/2021.05.31.446340)
- [2] Archie, E. A., Altmann, J. & Alberts, S. C. 2012 Social status predicts wound healing in wild baboons. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(23), 9017–9022. (doi:10.1073/pnas.1206391109)
- [3] Clauset, A., Newman, M. E. J. & Moore, C. 2004 Finding community structure in very large networks. *Physical Review E*, **70**(6), 066 111. (doi:10.1103/PhysRevE.70.066111)
- [4] Csardi, G. & Nepusz, T. 2006 The igraph software package for complex network research. *InterJournal*, **Complex Systems**, 1695.
- [5] DeDeo, S. & Hobson, E. A. 2021 From equality to hierarchy. (doi:10.1073/pnas.2106186118)
- [6] Feinerer, I., Hornik, K. & Meyer, D. 2008 Text mining infrastructure in r. *Journal of Statistical Software*, **25**(5), 1–54.
- [7] Fellows, I. 2018 *Package 'wordcloud'*. R package version 2.6.
- [8] Fitzpatrick, C. L., Hobson, E. A., Mendelson, T. C., Rodríguez, R. L., Safran, R. J., Scordato, E. S. C., Servedio, M. R., Stern, C. A., Symes, L. B. *et al.* 2018 Theory Meets Empiry: A Citation Network Analysis. *BioScience*, **68**(10), 805–812. (doi:10.1093/BIOSCI/BIY083)
- [9] Haddaway, N. R., Bethel, A., Dicks, L. V., Koricheva, J., Macura, B., Petrokofsky, G., Pullin, A. S., Savilaakso, S. & Stewart, G. B. 2020 Eight problems with literature reviews and how to fix them. *Nature Ecology & Evolution 2020 4:12*, **4**(12), 1582–1589. (doi:10.1038/s41559-020-01295-x)
- [10] Haddaway, N. R., Collins, A. M., Coughlin, D. & Kirk, S. 2015 The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. *PLOS ONE*, **10**(9), e0138 237. (doi:10.1371/JOURNAL.PONE.0138237)
- [11] Harzing, A.-W. 2013 *The Publish or Perish Book: Your guide to effective and responsible citation analysis*. Melbourne, Australia: Tarma Software Research, html edition edn.
- [12] Harzing, A.-W. & Alakangas, S. 2015 Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics 2015 106:2*, **106**(2), 787–804. (doi:10.1007/S11192-015-1798-9)
- [13] Hobson, E. A. 2020 Differences in social information are critical to understanding aggressive behavior in animal dominance hierarchies. *Current Opinion in Psychology*, **33**, 209–215. (doi:10.1016/j.copsyc.2019.09.010)
- [14] Hobson, E. A., Mønster, D. & DeDeo, S. 2021 Aggression heuristics underlie animal dominance hierarchies and provide evidence of group-level social information. *Proceedings of the National Academy of Sciences*, **118**(10).

- 353 [15] Kawakatsu, M., Chodrow, P. S., Eikmeier, N. & Larremore, D. B. 2021 Emergence of Hierarchy
354 in Networked Endorsement Dynamics. *PNAS*, **118**(16). (doi:10.1073/pnas.2015188118)
- 355 [16] Konno, K. 2020 Ignoring non-English-language studies may bias ecological meta-analyses.
356 *Ecol. Evol.*, **10**(13), 6373–6384. (doi:10.1002/ece3.6368)
- 357 [17] MacCormick, H. A., MacNulty, D. R., Bosacker, A. L., Lehman, C., Bailey, A., Anthony Collins,
358 D. & Packer, C. 2012 Male and female aggression: lessons from sex, rank, age, and injury in
359 olive baboons. *Behavioral Ecology*, **23**(3), 684–691. (doi:10.1093/beheco/ars021)
- 360 [18] McDonald, D. B. & Hobson, E. A. 2018 Edge weight variance: population genetic metrics for
361 social network analysis. *Animal Behaviour*, **136**, 239–250. (doi:10.1016/j.anbehav.2017.11.017)
- 362 [19] Nuñez, M. A. & Amano, T. 2021 Monolingual searches can limit and bias results in global
363 literature reviews. *Nature Ecology & Evolution 2021 5:3*, **5**(3), 264–264. (doi:10.1038/
364 s41559-020-01369-w)
- 365 [20] Sapolsky, R. M. 2004 Social Status and Health in Humans and Other Animals. *Annual Review*
366 *of Anthropology*, **33**(1), 393–418. (doi:10.1146/annurev.anthro.33.070203.144000)
- 367 [21] Sapolsky, R. M. 2005 The influence of social hierarchy on primate health. *Science (New York,*
368 *N. Y.)*, **308**(5722), 648–52. (doi:10.1126/science.1106477)
- 369 [22] Schjelderup-Ebbe, T. 1922 Beiträge zur Sozialpsychologie des Haushuhns (Contributions to
370 the social psychology of domestic chickens). *Zeitschrift für Psychologie*, **88**, 225–252.
- 371 [23] Shizuka, D. & McDonald, D. B. 2012 A social network perspective on measurements of dom-
372 inance hierarchies. *Animal Behaviour*, **83**(4), 925–934.
- 373 [24] Shizuka, D. & McDonald, D. B. 2015 The network motif architecture of dominance hierarchies.
374 *Journal of The Royal Society Interface*, **12**(105), 20150 080–20150 080.
- 375 [25] Silge, J. & Robinson, D. 2016 tidytext: Text mining and analysis using tidy data principles in r.
376 *JOSS*, **1**(3). (doi:10.21105/joss.00037)
- 377 [26] Strauss, E. D. & Holekamp, K. E. 2019 Social alliances improve rank and fitness in
378 convention-based societies. *Proceedings of the National Academy of Sciences of the United*
379 *States of America*, **116**(18), 8919–8924. (doi:10.1073/pnas.1810384116)
- 380 [27] Tarma Software Research Ltd 2021 Harzing’s Publish or Perish (version 7.31.3306.7768).
- 381 [28] Wickham, H. 2019 *stringr: Simple, consistent wrappers for common string operations*. R
382 package version 1.4.0.
- 383 [29] Youngblood, M., Baraghith, K. & Savage, P. E. 2020 Phylogenetic reconstruction of the cul-
384 tural evolution of electronic music via dynamic community detection (1975-1999).