



Correction

1

ARTICLE

DOI: 10.1057/s41599-018-0175-8

OPEN

A bibliometric analysis of the interdisciplinary field of cultural evolution

Mason Youngblood (1) 1,2 & David Lahti 1,2

ABSTRACT The science of cultural evolution is unified in its application of evolutionary logic to socially transmitted behavior, but diverse in methodologies and assumptions. Qualitative reviews have encouraged integration by illuminating points of divergence and fostering interaction. This effort would be greatly enhanced by quantitative data on patterns of collaboration and idea sharing in the literature. In the present study, we apply a novel combination of network, cluster, and bibliometric analyses to an extensive dataset of publications on cultural evolution, in order to represent the structure of the field and evaluate the level of disciplinary integration. We first construct a co-authorship network and identify subdisciplines. We then use bibliometric analyses to describe each subdiscipline and investigate trends in collaboration and productivity. Lastly, we assess the topographical distance and degree of citation sharing between subdisciplines, as well as the diversity of subject categories within subdisciplines. Our results reveal an increase in productivity and collaboration over time, albeit a higher inequality in author productivity than expected. Our structural approach reveals research subcommunities with differential levels of integration, citation sharing, and subject diversity. These findings confirm the emergence of a vigorous interdisciplinary field, and indicate ways to foster integration and synthesis in cultural evolution.

¹ Department of Psychology, The Graduate Center, City University of New York, New York, NY, USA. ² Department of Biology, Queens College, City University of New York, Flushing, NY, USA. Correspondence and requests for materials should be addressed to M.Y. (email: myoungblood@gradcenter.cuny.edu)

Introduction

he study of cultural evolution is an interdisciplinary endeavor that has captivated researchers from a diverse array of fields including evolutionary biology, anthropology, psychology, sociology, and computer science. This diversity of research communities and backgrounds has understandably been accompanied by a similarly broad range of approaches and methods (Richerson and Boyd, 2005; Driscoll, 2017; Mesoudi, 2015; Lewens, 2015). Evolutionary biology is often seen by cultural evolutionists as a reference point in terms of how an interdisciplinary science can be integrated upon a unified theoretical footing (Mesoudi et al. 2006). The history of evolutionary biology also illustrates, however, the diversity of working hypotheses or assumptions that can underlie research and reporting in an interdisciplinary field, often along disciplinary lines (Mayr, 1982). Classic typical differences of opinion between evolutionary paleontologists and neontologists, for instance, or between those who study change at the levels of molecules versus traits, have been covered in detail, and many of these differences are now reduced or even largely resolved (Adler et al. 2007; Hahn, 2008), although new or revised points of divergence are always arising (Laland et al. 2014; Pigliucci, 2006). Similarly, the various approaches of cultural evolutionists, especially in different disciplines, can carry with them different working hypotheses, or even explicit or implicit assumptions that are not shared across research communities. If such communities have little interaction or cooperation, these differences of opinion will become entrenched and hinder progress in the field. Understanding and synthesis among the various communities that perform cultural evolutionary research is currently underway, as represented for instance in the founding of the Cultural Evolution Society (CES). On the CES website, the participants of the 2015 workshop that inspired the founding of the society are described as coming from a "melting pot of disciplines that need to become integrated to create a science of cultural change informed by evolutionary theory¹". The integration of this "melting pot" could help to recontextualize studies from different disciplines to strengthen the field's empirical base (Mesoudi et al. 2006; Henrich et al. 2008), and to increase the application of cultural evolutionary knowledge to policy matters (Gibson and Lawson, 2015). Additionally, as the application of evolutionary logic becomes more popular in the public sphere, the field of cultural evolution will benefit by fostering a stronger, more unified empirical and methodological foundation for evolutionary approaches to sociocultural questions (Smith et al. 2001; Mace, 2014).

Proposals to integrate knowledge from diverse disciplines often involve the encouragement of interdisciplinary research. Interdisciplinary research is becoming more common (Porter and Rafols, 2009) and increasingly en vogue, as evidenced by the fivefold increase in publications tagged as {"interdisciplinary"} on Web of Science between 2002 and 2016, which is twice the increase in Web of Science entries as a whole over that period. In fact, both the U. S. National Science Foundation and National Institutes of Health have made the support of interdisciplinary research an explicit goal of their respective strategic initiatives (National Science Foundation, 2014; National Institutes of Health, 2016). Although interdisciplinarity is often used as a rallying cry among scientists, some have advised caution (Jacobs and Frickel, 2009). A recent study found that researchers collaborating across disciplinary boundaries tend to publish fewer, more highly cited papers, but, interestingly, this "productivity penalty" was significantly lower for authors in more interdisciplinary fields (Leahey et al. 2017). This indicates that, for authors in an already interdisciplinary field such as cultural evolution, the benefits of increased citation probably outweigh

any costs to productivity. Additionally, the critique that interdisciplinarity could divert focus from more targeted, mechanistic questions assumes that fields have already developed the methodological underpinnings to investigate those questions. For a less developed field like cultural evolution, in which researchers have been working with different assumptions for decades (Acerbi and Mesoudi, 2015; Claidière and André, 2011), theoretical unification could lead to an increase in empirical investigation and "more dynamic applications of emerging knowledge" (Aboelela et al. 2007).

In order to understand the interdisciplinarity of cultural evolution, either with respect to the diversity of thought and practice or to the degree to which research communities communicate with each other, we have to know who those research communities are. Previous suggestions have depended upon broad, qualitative and subjective descriptions of field structure (Wimsatt, 2013; Mesoudi et al. 2006; Henrich et al. 2008). These are valuable, especially as they highlight frequent points of divergence in ideas, and promote simple conceptualization of possible challenges. However, only quantitative data on the structure of the field of cultural evolution can delineate actual communities, determine whether they fall along disciplinary lines, assess degrees of communication and cross-fertilization, and inform specific goals about how to encourage integration and synthesis. Quantitative methods can also provide a baseline for future studies to determine how well such synthesis is proceeding.

Bibliometrics, or the statistical analysis of published works, was first described by Pritchard (1969), and bibliometric approaches have since been used to analyze scientific literature in fields as diverse as neuroscience (Yeung et al. 2017), water security (Huai and Chai, 2016), and bariatric surgery (Dabi et al. 2016). Additionally, bibliometric methods have been used to support historical assumptions about the development of fields (Raina and Gupta, 1998), identify connections between scientific growth and policy changes (das Neves Machado et al. 2015), and investigate the collaborative structure of an interdisciplinary field (Liu and Xia, 2015). Evidence suggests that collaboration may positively increase productivity (Landry et al. 1996; Lee and Bozeman, 2005), impact (Larivière et al. 2014), and citation rate (Figg et al. 2006), making it a key feature of bibliometric studies. Historically, bibliometric analyses have been limited by data accessibility, but the establishment of web-based citation indexing databases has increased their scope. Web of Science (WoS), a database with records from more than 33,000 journals (Clarivate Analytics, 2017), is a particularly useful resource for bibliometrics research because each record is tagged with key information such as author name, source, cited references, keywords, and research area. Additionally, users are able to download the complete metadata for every publication resulting from a given search term.

Access to these nearly exhaustive bibliometric datasets, in combination with social network analysis, has allowed researchers to investigate previously inaccessible questions about how scientific fields are structured. The construction of co-authorship networks, in which each node represents an author and each link represents a co-authorship relationship, can provide particularly valuable insight into the topography of a given field. Coauthorship represents a direct working relationship between authors, and is often used as a proxy measure for scientific collaboration in bibliometric studies. Unfortunately, large networks are challenging to visualize with traditional mapping methods. VOS (visualization of similarity) mapping is a recently developed technique that allows representative and intelligible visualization of more complex networks (van Eck and Waltman, 2007), and has been combined with modularity-based clustering to parse networks into discrete clusters based on the association strength

among nodes (Waltman et al. 2010). We utilize VOS mapping in this study because of its ability to simultaneously map and identify clusters within a network, and its widespread use in recent bibliometric analyses (Taşkin and Aydinoglu, 2015; Sweileh et al. 2016a, 2016b).

One challenge in using network analysis to evaluate the structure of a field is the accurate identification and labeling of subdisciplines. Previous studies have largely depended upon the subject categories developed by the Institute for Scientific Information (ISI) and assigned by WoS, but these categories are general, retrospective, and do not adequately describe the structures of emerging fields such as cultural evolution (Pudovkin and Garfield, 2002; Wagner et al. 2011). Rather than sorting authors into existing scientific categories, we chose to take a "structural approach" by classifying subdisciplines a posteriori, based on the discrete clustering patterns identified in the co-authorship network during the mapping process (Wagner et al. 2011). One of the more recently developed methods for classifying subdisciplines involves the identification and description of core documents in co-citation and bibliographic coupling networks (Glänzel and Thijs, 2011, 2017). Glänzel and Thijs (2011) found this method to be particularly effective for identifying topics in emerging fields, where text-based classification (i.e., keyword usage) might overestimate the degree of overlap between subdisciplines. We chose to apply this method to our co-authorship network by identifying a set of core authors for each cluster (defined by the number of co-authorship links they have with other nodes) and evaluating their disciplinary affiliations, methodological approaches, and intellectual contributions. Although this approach increases the difficulty of placing authors into discrete disciplinary categories, it allows for more realistic and nuanced classification of clusters.

By employing a novel combination of network, cluster, and bibliometric analyses to an extensive dataset of journal article metadata, we aim to determine the degree of intellectual integration among subdisciplines in the field of cultural evolution, in order to inform more targeted efforts towards future synthesis.

Methods

All articles used in this study were retrieved from the WoS Core Collection. The search term {"cultural evolution"} was used in the topic field, and results were filtered by publication year through 2017. The complete metadata for each resulting publication, as well as a subject category frequency list, was compiled and manually exported on July 10, 2018.

VOSviewer (version 1.6.8), a network analysis software tool used to analyze bibliometric data, was used to construct a collaboration network based on co-authorship (van Eck and Waltman, 2009). In order to prevent articles with many co-authors from biasing network structure, we utilized fractional counting and excluded articles with more than 10 co-authors (Perianes-Rodriguez et al. 2016). The resolution was set to 0.03, the highest setting that resulted in a manageable number of clusters with suitable sample sizes, with no minimum cluster size. Any redundant names were eliminated using a thesaurus file, so that each author with multiple entries was designated by their last name and first initial (i.e., RICHERSON PJ to RICHERSON P). In order to avoid author name disambiguation, which is known to influence the structure of co-authorship networks, individuals with the same last name and first initial were removed from the thesaurus file (Barbastefano et al. 2015). To qualitatively describe each cluster we designated the five authors with the most coauthorship links as "core authors", and evaluated their departmental affiliations and experimental approaches. For additional support, we identified country affiliations of first authors, authorchosen keywords, and subject categories that were statistically over-represented in each cluster using GeneMerge, a software package for identifying categorical variables that occur at higher frequencies in data subsets (Castillo-Davis and Hartl, 2003).

The subject category frequency list, which lists of the number of publications in the dataset that fall within each subject category assigned by WoS (i.e., "anthropology", "behavioral sciences", "evolutionary biology"), was used to generate a science overlay map in Pajek (Rafols et al. 2010). Science overlay maps superimpose data from a local network over a "global map of science" based on a co-citation matrix of ISI subject categories from 2007, providing an overview of how the focal field is situated in the rest of science (Rafols et al. 2010). Additional science overlay maps were constructed from the subject categories used within each cluster of co-authors identified in VOSviewer. The subject category frequency list was also used to calculate the Stirling-Rao diversity measure, an indicator of interdisciplinarity that takes into account the distances between subject categories in the network, for both the entire network and each cluster (Rao, 1980; Stirling, 2007; Leydesdorff and Rafols, 2011). Further details about science overlay maps and the Stirling-Rao diversity measure are available online².

CINNA (version 1.1.14), an R package for network analysis, was used to identify and calculate the two traditional centrality measures that explained the most variance in network structure: eigenvector and closeness centrality (Ashtiani et al. 2017). Eigenvector centrality is a weighted measure of the number of links that a given node in the network has with other nodes, while closeness centrality indicates the length of the most efficient paths to all other nodes in the network. These measures act as proxies for influence and distance from the center, respectively (Landherr et al. 2010). Both measures were rescaled to fall between 0 and 1. By comparing the centrality values for nodes in each cluster, we can quantitatively characterize the topography of the coauthorship network.

The R packages qgraph (version 1.4.4) and igraph (version 1.1.2) were used to calculate the small-worldness index and network density, respectively. The small-worldness index indicates whether or not a network exhibits the small-world structure shown by many collaboration networks with high clustering and short average path lengths (Humphries and Gurney, 2008; Ebadi and Schiffauerova, 2015a, 2015b). Values greater than one reflect small-world structure. Network density is the ratio of observed edges to possible edges in a network.

Bibliometrix (version 1.7), an R package for bibliometric analysis, was used to import and manage the metadata from WoS. ³Articles published by individuals in each group were sorted accordingly. Any article co-authored with a member of another group was excluded to allow for more parsimonious subdisciplinary classification, and to avoid exaggerating the degree of citation overlap between groups. The exclusion of overlapping articles occurred downstream of the network analysis, and only impacted the group-level descriptive data and citation analyses. Basic descriptive data was generated for each group, such as authors, sources, keywords, and subjects. Each author's h-index, a common proxy measure for individual scientific output, was calculated using only the records in the dataset to determine individual impact within the field (Hirsch, 2005). The annual collaboration index, a ratio of the number of authors of coauthored articles to the total number of co-authored articles, was calculated for the entire dataset and each cluster. In addition, the Lotka's law coefficients were estimated to determine the relative frequency of authors ranked by productivity (Lotka, 1926). Lotka's law predicts that the relative frequency distribution of author productivity for any given field will be a hyperbolic, inverse

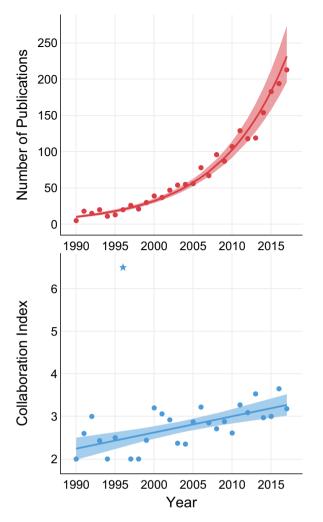


Fig. 1 Scatterplots showing the number of articles published on the topic of cultural evolution (top), as well as the collaboration index of the field (bottom), each year. Exponential fitting of the data revealed an increase in the number of articles written between 1990 and 2017 ($r^2 = 0.9443$). Linear fitting of the data, after the removal of the outlier indicated by the star, revealed an increase in the collaboration index as well ($r^2 = 0.4333$)

square function ($y = 1/x^2$), such that a minority of the authors in the dataset are publishing the majority of the articles. The deviation of the observed function from the predicted inverse square function acts as a metric for the inequality in productivity of the field. Lotka's law breaks down when fractional counting is used and/or the number of collaborators on any article in the dataset is extremely high (Rousseau, 1992; Kretschmer and Rousseau, 2001), so we ensured that full counting was used and none of the articles in the dataset exceeded 100 co-authors.

In order to determine how much information was shared among the co-authorship groups, the collective references cited by authors in each group were compiled and trimmed to eliminate duplicates. The overlaps between the groups' collective reference lists were analyzed using SuperExactTest (version 0.99.4), an R package developed to statistically evaluate multi-set intersections by calculating the probability of set overlaps given the size of the entire dataset (Wang et al. 2015). Lastly, a word-cloud was constructed in R to visually compare author-chosen keyword usage among the five groups. All analyses in R were conducted using version 3.3.3.

Results

The search criteria yielded 2091 records published by 3451 authors. Sixty-seven countries are represented in the dataset, but the majority of articles were published by first authors in the United States (34.78%), England (12.31%), Canada (5.71%), Germany (5.06%), and Scotland (4.75%). Although 685 journals are represented in the dataset, the five most productive were PNAS (3.35%), Behavioral and Brain Sciences (3.29%), Philosophical Transactions of the Royal Society B (2.93%), PLoS ONE (2.46%), and Evolution and Human Behavior (2.25%). The overall collaboration index (CI), or the average number of authors on multi-authored papers, was 2.43. The small-worldness index, calculated against 1000 random networks, was 22.42, indicating that the network exhibits small-world structure. Other descriptive network statistics can be found in Table S1. Temporal analyses indicate that the field as a whole has become more productive over time, with an annual percentage growth rate of 13.61%. Exponential fitting of the data revealed an increase in the number of articles written between 1990 and 2017 ($r^2 = 0.9443$) (top panel of Fig. 1). Upon first analysis, a linear increase in the CI across the same time period was not evident, but the removal of an outlier identified by its studentized residual (year = 1996, CI = 6.5; p < 0.001) revealed a positive relationship ($r^2 = 0.4333$) (bottom panel of Fig. 1). Linear fitting also revealed an increase in the mean number of unique countries associated with each article written between 1990 and 2017 ($r^2 = 0.5966$).

The Lotka's law estimation yielded a beta coefficient (β) of 2.43 and a constant coefficient (c) of 0.40 ($r^2=0.9430$). In the entire dataset, 2907 authors (84.24%) published a single paper, 312 authors (9.04%) published two papers, and 232 authors (6.72%) published three or more papers. The most productive author, Alex Mesoudi, published 35 (1.67%) of the papers included in the dataset. The Kolmogorov–Smirnov test found the Lotka's law estimation to be significantly different from the theoretically predicted inverse square function (p < 0.05), indicating that the relative frequency distribution of the authors' number of publications in the dataset is more skewed right than predicted (see Fig. S2). This suggests that the field of cultural evolution does not conform to Lotka's law, and has fewer highly productive authors than expected.

The largest set of authors linked by co-authorship (n = 629)was used for the network analysis, resulting in the seven clusters portrayed in Fig. 2 ($n_1 = 183$; $n_2 = 146$; $n_3 = 134$; $n_4 = 75$; $n_5 = 188$ 56; $n_6 = 20$; $n_7 = 15$). Articles published by individuals in each group, excluding those co-authored with members of other groups, were compiled in R $(n_1 = 177; n_2 = 131; n_3 = 137; n_4 =$ 65; $n_5 = 57$; $n_6 = 6$; $n_7 = 9$). The mean h-indices of the authors in each group were calculated using only articles in the dataset across all years to examine collective output and impact (M_1 = 2.04; $M_2 = 2.30$; $M_3 = 1.82$; $M_4 = 2.06$; $M_5 = 2.14$; $M_6 = 1.1$; $M_7 = 2.06$; $M_8 = 2.06$; $M_9 = 2.0$ = 1.93). A one-way ANOVA found that the differences between the mean h-indices of each group were non-significant but trending (p = 0.069). Pairwise comparisons using t-tests with pooled standard deviations indicated that the mean h-index of group 6 is significantly lower than the mean h-indices of all groups except for group 3 (p < 0.05). The CI for each group was calculated ($CI_1 = 1.79$; $CI_2 = 1.78$; $CI_3 = 1.48$; $CI_4 = 1.81$; $CI_5 = 1.48$; $CI_{10} = 1.81$; $CI_{10} = 1.8$ 1.4; $CI_6 = 10.3$; $CI_7 = 2.11$). Even after the removal of an extreme outlier (42 authors) the CI of group 6 ($CI_6 = 4$) was nearly twice that of group 7. This indicates that group 6 is the most internally collaborative in the dataset.

The descriptive statistics of the core authors, and the statistically over-represented country affiliations of first authors, author-chosen keywords, and subject categories in each cluster, can be found in Table S3 and Table S4, respectively. A wordcloud of author-chosen keywords can be found in Fig. S5. Core authors in

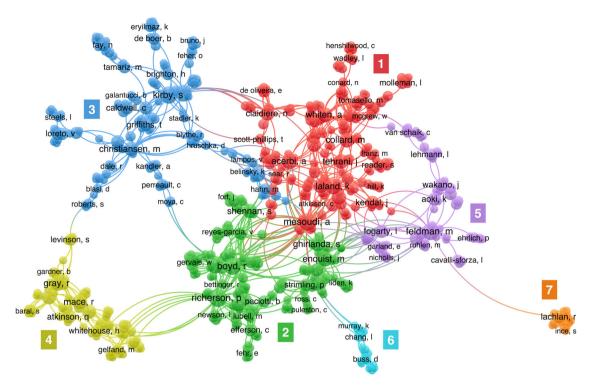


Fig. 2 The largest connected co-authorship network in the dataset, analyzed using VOS clustering (n = 629). Red corresponds to group 1 ("biological anthropology and archeology"; $n_1 = 183$), green corresponds to group 2 ("mathematical modeling and dual-inheritance theory"; $n_2 = 146$), blue corresponds to group 3 ("cognitive linguistics and experimental cultural evolution"; $n_3 = 134$), yellow corresponds to group 4 ("cross-cultural and phylogenetic studies"; $n_4 = 75$), purple corresponds to group 5 ("computational biology and cultural niche construction"; $n_5 = 56$), cyan corresponds to group 6 ("evolutionary psychology"; $n_6 = 20$), and orange corresponds to group 7 ("behavioral ecology and birdsong"; $n_7 = 15$). Name size indicates total link strength. Many authors were arbitrarily excluded from the figure by the visualization algorithm in VOSviewer to maximize legibility. A complete, interactive version of the network can be found in the Dataverse repository entry: https://doi.org/10.7910/DVN/LBIDEL

group 1 included Alex Mesoudi, Alberto Acerbi, Kevin Laland, Mark Collard, and Andrew Whiten. The keywords "social learning", "conformity", and "chimpanzees", as well as the subject categories "anthropology" and "biological psychology" are overrepresented. We have chosen to label this group (1) biological anthropology and archeology. Core authors in group 2 included Peter Richerson, Robert Boyd, Joseph Henrich, Magnus Enquist, and Stefano Ghirlanda. The keywords "cooperation", "religion", and "cultural transmission", as well as the subject category "behavioral sciences" are over-represented. We have chosen to label this group (2) mathematical modeling and dual-inheritance theory. Core authors in group 3 included Simon Kirby, Kenny Smith, Alex Bentley, Morten Christiansen, and Christine Caldwell. The keywords "language evolution", "iterated learning", and "compositionality", as well as the subject categories "experimental psychology" and "linguistics" are over-represented. We have chosen to label this group (3) cognitive linguistics and experimental cultural evolution. Core authors in group 4 included Ruth Mace, Russell Gray, Simon Greenhill, Thomas Currie, and Quentin Atkinson. The keywords "cultural phylogenetics", "phylogenetic comparative methods", and "cross-cultural research", as well as the subject category "biology" are overrepresented. We have chosen to label this group (4) cross-cultural and phylogenetic studies. Core authors in group 5 included Marcus Feldman, Nicole Creanza, Kenichi Aoki, Laurel Fogarty, and Joe Wakano. The keywords "demography", "cultural accumulation", and "gene-culture coevolution", as well as the subject categories "mathematical and computational biology" and 'genetics and heredity" are over-represented. We have chosen to label this group (5) computational biology and cultural niche construction. Core authors in group 6 included David Buss, Lei Chang, Hui Jing Lu, Todd Shackelford, and Laith Al-Shawaf. The keywords "mate preferences", "evolutionary psychology", and "sex", as well as the subject category "social psychology" are over-represented. We have chosen to label this group (6) evolutionary psychology. Core authors in group 7 included Robert Lachlan, Peter Slater, Maria Servedio, Carel ten Cate, and Machteld Verzijden. The keywords "song learning", "bird song", and "speciation", as well as the subject categories "ecology" and "evolutionary biology" are over-represented. We have chosen to label this group (7) behavioral ecology and birdsong. The subdisciplinary labels that we have applied to each group of co-authors are ultimately subjective and should be treated as conceptual guides rather than discrete categories.

Science overlay maps were constructed using the subject category frequency lists for the entire network, as well as for each cluster of co-authors identified in VOSviewer (see Fig. S6). All 19 of the broad categories identified as explaining the majority of the variance in the global science map were present in the full network (Leydesdorff et al. 2012), but they were represented to different extents in each of the five groups. The Stirling-Rao diversity measure for each group was calculated as well (D_1 = 0.7196; D_2 = 0.7467; D_3 = 0.7357; D_4 = 0.6660; D_5 = 0.5354; D_6 = 0.5768; D_7 = 0.4502). This indicates that group 2 is the most interdisciplinary in the dataset.

Based on the eigenvector centrality values, group 2 has the highest weighted connectivity with other nodes in the network ($M_1=0.0029;\ M_2=0.0376;\ M_3=0.0011;\ M_4=0.0026;\ M_5=0.0052;\ M_6=0.0004;\ M_7=0.0006)$ (see Fig. 3). A one-way ANOVA found significant differences between the mean eigenvector centrality values of each group (p<0.001). Pairwise comparisons using t-tests with pooled standard deviations indicated

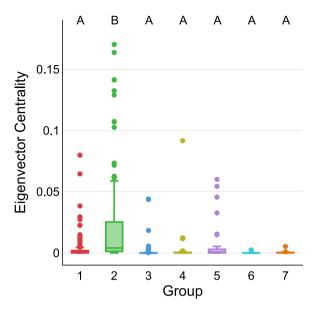


Fig. 3 The distribution of eigenvector centrality values for the authors in each group. The three most extreme outliers in group 2 were excluded from the figure for higher resolution. Pairwise comparisons were conducted using *t*-tests with pooled standard deviations, and clusters with significantly different means are denoted with different letters

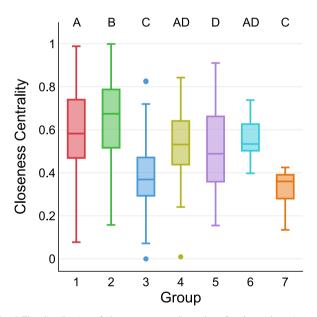


Fig. 4 The distribution of closeness centrality values for the authors in each cluster. Pairwise comparisons were conducted using t-tests with pooled standard deviations, and clusters with significantly different means are denoted with different letters

that the mean of group 2 is significantly higher than that of all other groups (p < 0.05), while the means of all other groups are not significantly different from one another (p > 0.05).

Based on the closeness centrality values, group 2 is the closest to the center of the network ($M_1 = 0.5870$; $M_2 = 0.6554$; $M_3 = 0.3876$; $M_4 = 0.5488$; $M_5 = 0.5134$; $M_6 = 0.5562$; $M_7 = 0.3299$) (see Fig. 4). A one-way ANOVA found significant differences between the mean closeness centrality values of each group (p < 0.001). Pairwise comparisons using t-tests with pooled standard deviations indicated that the mean of group 2 is significantly higher than those of all other groups (p < 0.05). Additional

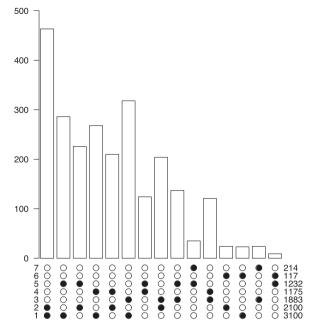


Fig. 5 The degree of citation overlap between groups. The groups included in each pairwise comparison are designated by the black dots below each bar. For each pairwise comparison, the *y*-axis is the number of articles that are shared between the two collective reference lists. The total size of each collective reference list appears to the right of each row of dots. Only pairwise comparisons with statistically significant intersections are included, and the results are sorted by increasing *p*-value

pairwise comparisons indicated that the means of groups 3 and 7 are significantly lower than those of all other groups (p < 0.001), but are not significantly different from one another (p > 0.05).

The degree of citation overlap between groups is shown in Fig. 5. Due to inconsistencies in citation formatting and completeness, only references with a digital object identifier (DOI) were included in the analysis (41.5%). Groups 1 and 2, and groups 1 and 5, have the highest degree of citation overlap, and the results of SuperExactTest indicate that 71.4% of intersections between pairs of groups are significantly higher than chance given the size of the entire dataset (p < 0.05).

Discussion

The study of cultural evolution is of burgeoning interest, with a publication rate rising twice as quickly as science in general. Over the last 25 years, the number of professional journal articles covering cultural evolution has increased dramatically and more countries are regularly involved in publishing. The increase in the collaboration index is particularly promising, given the evidence that larger research teams have higher scientific impact (Larivière et al. 2014) and productivity (Landry et al. 1996; Lee and Bozeman, 2005). As cultural evolution draws researchers and ideas from a broad spectrum of academic fields, the present review has focused especially on assessing interdisciplinary interaction and integration. Despite vigorous research activity, the field's youth is indicated by the absence of a professional society until 2015, and the continued absence of a dedicated academic journal. Cultural evolutionary research is published in an unusually thinly distributed manner across publishing outlets. Authorship is disparate, with most authors publishing only a single study, and fewer highly productive authors in the field than expected. Collaborations coalesce within seven topical clusters that differ in their level of interaction within and between groups, although the clusters overlap substantially in the references they cite.

Vocabularies likewise have common elements, but each cluster employs distinctive terms as well. In general the analysis paints a picture of an active and energetic field that is increasing in integration, but that as yet has only partially established the communicative and collaborative networks typical of a cohesive field of study.

The estimation of Lotka's law for this dataset indicated significantly more disparity in author productivity than expected. A recent productivity analysis of 30 scientific fields found the average proportion of authors with one publication to be 69% (Ruiz-Castillo and Costas, 2014), which is approximately 15% lower than our result for cultural evolution. This high inequality of productivity in cultural evolution could be a result of a tendency for most researchers to be only occasionally interdisciplinary. As researchers from various fields begin to publish on the same topic, probably only a small minority of them will be consistently interdisciplinary enough to have high productivity related to the emerging topic, whereas the majority of researchers will continue to publish primarily in their home disciplines.

Major publishing outlets for the field have not yet emerged. According to data from ISI Web of Science, the journal with the highest use by cultural evolutionists (PNAS) accounts for only 3.35% of the publications, and the top five together (the only journals with >2%) sum to only 14.28%. By contrast, the single journals publishing the most studies in several other active topic areas, for instance molecular evolution, the evolution of cooperation, and sexual selection, account for over 7% of each subfield's produce by the same metrics, and the top 5 journals in these areas publish over 21-27%, with 28-40% of studies accounted for by 8–10 journals that each publish >2% of the field. The reason why dozens of journals are required to find a similar proportion of cultural evolution's output is not clear. A dedicated journal would probably help this situation, as they cover 7-8% of the field in many cases, but this does not always happen. Neither evolutionary psychology nor invasive species have a robust core group of journals, for instance, despite having dedicated journals and much research activity overall. Coevolution and niche construction are two additional topic areas with focused research activity but a lack of consistent publication outlets, although cultural evolution is the most thinly distributed of all of these

Our network analysis yielded seven clusters based on coauthorship patterns, meaning that researchers tend to collaborate with others within particular subgroups in the field. Our structural approach to mapping the field means that instead of determining group membership a priori (e.g., by departmental affiliations), we allowed the behavior of researchers (coauthorships) to speak for itself. However, this approach raises the challenge of labeling the resulting subdisciplines. The full diversity of approaches within any cluster cannot be exhaustively described without overestimating the degree of overlap between clusters. To avoid this, we chose to classify subdisciplines by characterizing the contributions of the core authors within them. Thus the labels that we proposed in the results should be taken as subjective, flexible indicators of researcher communities rather than strict subdisciplinary definitions.

The seven co-authorship clusters differ in their activity and level of integration. "Mathematical modeling and dual-inheritance theory" has the greatest weighted connectivity to other nodes, closeness to the center of the network, and subject category diversity, indicating that it is the most influential, central, and intellectually broad cluster. This is unsurprising given the fact that several of its members played integral roles in founding the field, alongside members of "computational biology and cultural niche construction" (Mesoudi, 2015). "Cognitive

linguistics and experimental cultural evolution" and "behavioral ecology and birdsong" are furthest from the center of the network. "Biological anthropology and archeology" and "mathematical modeling and dual-inheritance theory", as well as "biological anthropology and archeology" and "computational biology and cultural niche construction", overlap most in cited references, but over 70% of group pairs have more citation overlap than expected. All insignificant citation overlaps are between either "evolutionary psychology" or "behavioral ecology and birdsong", the two smallest groups identified in the field, and non-adjacent groups. This indicates that intellectual similarity is related to structural distance in the collaboration network. "Evolutionary psychology" has the lowest mean h-index and is the most internally collaborative, which could be a result of its small size compared to other subdisciplines.

Recent research has demonstrated that the relationship between interdisciplinary collaboration and impact scales up with disciplinary distance (Larivière et al. 2015), indicating that researchers in the field of cultural evolution could reap even more benefits from reaching out to authors further away in the network. Based on the results of this study, increased collaboration between the two most peripheral groups, "cognitive linguistics and experimental cultural evolution" and "behavioral ecology and birdsong", and the rest of the network would have the greatest impact on the integration of the field. A recent review of interdisciplinary research in the field actually highlighted the value of both cognitive and animal behavior approaches in connecting micro-level processes to macro-level patterns in cultural evolution (Kolodny et al. 2018). "Mathematical modeling and dualinheritance theory" may be the best candidate for collaborations with "cognitive linguistics and experimental cultural evolution", as it is highly connected to the rest of the network and there is already a great deal of intellectual overlap between these two groups. "Computational biology and cultural niche construction" is probably the best candidate for collaborations with "behavioral ecology and birdsong" as they already share co-authorship connections, but "behavioral ecology and birdsong" actually has a much older average publication year. More recent researchers who utilize approaches from behavioral ecology (i.e., Lucy Aplin and Johan Lind) already appear to be more collaborative with other groups, so the apparent distance between this subdiscipline and the rest of the network could be due to historical lag.

Recent studies in cultural evolution have shown that partial connectivity may enhance cultural complexity over full connectivity, using small experimental groups (Derex and Boyd, 2016) and models incorporating population size and fragmentation (Derex et al. 2018). Other studies in the innovation and management literature have found that small-world networks, characterized by higher clustering and shorter path lengths, have higher innovation rates and reach solutions more quickly (Schilling and Phelps, 2007; Mason et al. 2005). The apparent discrepancy in these results could be because collaboration networks often exhibit small-world structure (Ebadi and Schiffauerova, 2015a, 2015b) and are thus below the threshold of partial connectivity used in the cultural evolution studies. Given the co-authorship network's small-world structure and low density, efforts to integrate the field are more likely to enhance innovation through increased information access than to impair it through full connectivity.

Several limitations of this study are worth mentioning. First, the search term {"cultural evolution"} excludes work by researchers who investigate changes in cultural traits but use different terminology. Although including other terms such as "cultural change" and "language evolution" would have expanded our results, we chose to limit our search to one term to avoid

over-representing particular subdisciplines. Second, many of the most influential books and book chapters on cultural evolution are not indexed on WoS. As a result, fields with more emphasis on long-form publishing, such as cultural anthropology, may be underrepresented in this analysis. Third, our analysis excludes single-author papers, as we focused on coauthorship as an indicator of collaboration and field integration. Fourth, earlier articles are less likely than more recent articles to be represented in the coauthorship network because the authors of older papers are less likely to be collaborating recently. This bias actually aided us in avoiding another bias, however: the term "cultural evolution" was occasionally used, especially before 1975, in a way inconsistent with contemporary usage, specifically reflecting a now-outmoded unilineal approach to cultural change and development. Few if any of the papers in our dataset used the term in this context. Lastly, our group-level results depend upon the resolution settings of the cluster analysis. Although larger resolutions break down the clusters into smaller subcommunities, they also reduce sample sizes and make classification less manageable. Increasing our resolution by 0.01, the smallest possible unit, resulted in 10 clusters with sample sizes down to 10 authors, so we used 0.03 to maximize our statistical power and classification ability.

Cultural evolution as a whole draws together theory and data from a remarkable variety of sources, and does so by an equally broad array of methods (Kolodny et al. 2018). The field has already demonstrated that this intellectual diversity, bound together with a broad evolutionary perspective, will enable us to understand and explain change in socially learned traits over time in a new and powerful way (Mesoudi, 2017). However, wherever there are groups that have differences of approach, differences in vocabulary, and differences in researchers with more collaboration within than between groups, we can expect to find differences of assumptions and possibly conclusions as well. Thus, the study of cultural evolution needs more integration if it is to function as a scientific community. Bridging efforts could take several forms. Workshops geared towards interdisciplinary research have been shown to improve integration in other fields (Piso et al. 2016; Barrows et al. 2008), and could be easily integrated into the annual meetings of the Cultural Evolution Society. Another potential strategy is to implement a database for researchers to find collaborators with complementary interests and skill sets (Novak et al. 2014), to ensure that co-authors are enhancing creativity by bringing new knowledge to the team (Lee et al. 2015). In addition, strategic grant funding targeted at integrative research could begin to reduce the distance between subdisciplines. Recent evidence that impact per dollar is actually higher for smaller grants (Fortin and Currie, 2013) suggests that a more distributed funding format might be more effective for increasing the visibility of the field as a whole. Since younger scientists tend to be more collaborative and act as intellectual bridges between disciplines (Ebadi and Schiffauerova, 2015a, 2015b), providing more opportunities for early-career grant funding could also increase the connectivity of the entire network. Lastly, fostering a culture of open science, in which researchers publish and exchange their data and scripts, could play a key role in promoting interdisciplinarity and trust in a field with such an emphasis on mathematical modeling and phylogenetic methods (Mesoudi, 2017; Easterbrook, 2014). By achieving conceptual synthesis through coordinated, interdisciplinary collaboration, the field of cultural evolution will be better equipped to tackle the "grand challenges" it faces (Brewer et al. 2017).

Received: 8 June 2018 Accepted: 20 September 2018 Published online: 02 October 2018

Notes

- 1 The Evolution Institute: http://bit.ly/2tmeJZ6
- 2 A user-friendly method for generating overlay maps (2012 update): https://bit.ly/ 2mmTOF
- 3 Bibliometrix: http://bit.ly/2t21Dkk

References

- Aboelela SW, Larson E, Bakken S, Carrasquillo O, Formicola A, Glied SA, Haas J, Gebbie KM (2007) Defining interdisciplinary research: conclusions from a critical review of the literature. Health Serv Res 42(1p1):329–346
- Acerbi A, Mesoudi A (2015) If we are all cultural Darwinians what's the fuss about? Clarifying recent disagreements in the field of cultural evolution. Biol Philos 30(4):481–503
- Adler PB, HilleRisLambers J, Levine JM (2007) A niche for neutrality. Ecol Lett 10 (2):95–104
- Ashtiani M, Mirzaie M, Jafari M (2017) CINNA: an R package for deciphering central informative nodes in network analysis. bioRxiv 168757
- Barbastefano RG, Souza C, de Sousa Costa JM, Teixeira PM (2015) Influência da ambiguidade de nomes na centralidade de redes de coautoria. Transinformação 27(3):189-198
- Barrows FT, Bellis D, Krogdahl Å, Silverstein JT, Herman EM, Sealey WM, Rust MB, Gatlin DM (2008) Report of the plant products in aquafeed strategic planning workshop: An integrated interdisciplinary research roadmap for increasing utilization of plant feedstuffs in diets for carnivorous fish. Rev Fish Sci 16(4):449–455
- Brewer J, Gelfand M, Jackson JC, MacDonald IF, Peregrine PN, Richerson PJ, Turchin P, Whitehouse H, Wilson DS (2017) Grand challenges for the study of cultural evolution. Nat Ecol Evol 1(3):70
- Castillo-Davis CI, Hartl DL (2003) GeneMerge-post-genomic analysis, data mining, and hypothesis testing. Bioinformatics 19(7):891–892
- Claidière N, André J-B (2011) The transmission of genes and culture: a questionable analogy. Evol Biol 39(1):12–24
- Clarivate Analytics (2017) Web of Science Fact Book. Technical report
- Dabi Y, Darrigues L, Katsahian S, Azoulay D, Antonio MD, Lazzati A (2016) Publication trends in bariatric surgery: a bibliometric study. Obes Surg 26 (11):2691–2699
- das Neves Machado R, Vargas-Quesada B, Leta J (2015) Intellectual structure in stem cell research: exploring Brazilian scientific articles from 2001 to 2010. Scientometrics 106(2):525–537
- Derex M, Boyd R (2016) Partial connectivity increases cultural accumulation within groups. PNAS 113(11):2982–2987
- Derex M, Perreault C, Boyd R (2018) Divide and conquer: intermediate levels of population fragmentation maximize cultural accumulation. Philos Trans R Soc B 373:20170062
- Driscoll C (2017) The evolutionary culture concepts. Philos Sci 84(1):35-55
- Easterbrook SM (2014) Open code for open science? Nat Geosci 7(11):779-781
- Ebadi A, Schiffauerova A (2015a) How to become an important player in scientific collaboration networks? J Informetr 9(4):809–825
- Ebadi A, Schiffauerova A (2015b) On the relation between the small world structure and scientific activities. PLoS One 10(3):e0121129
- Figg WD, Dunn L, Liewehr DJ, Steinberg SM, Thurman PW, Barrett JC, Birkinshaw J (2006) Scientific collaboration results in higher citation rates of published articles. Pharmacotherapy 26(6):759–767
- Fortin J-M, Currie DJ (2013) Big science vs. little science: how scientific impact scales with funding. PLoS One 8(6):e65263
- Gibson MA, Lawson DW (2015) Applying evolutionary anthropology. Evolut Anthropol 24(1):3–14
- Glänzel W, Thijs B (2011) Using'core documents' for detecting and labelling new emerging topics. Scientometrics 91(2):399–416
- Glänzel W, Thijs B (2017) Using hybrid methods and core documents' for the representation of clusters and topics: the astronomy dataset. Scientometrics 111(2):1071–1087
- Hahn MW (2008) Toward a selection theory of molecular evolution. Evolution 62 (2):255–265
- Henrich J, Boyd R, Richerson PJ (2008) Five misunderstandings about cultural evolution, human. Nature 19(2):119–137
- Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Natl Acad Sci 102(46):16569–16572
- Huai C, Chai L (2016) A bibliometric analysis on the performance and underlying dynamic patterns of water security research. Scientometrics 108 (3):1531–1551
- Humphries MD, Gurney K (2008) Network 'Small-World-Ness': A quantitative method for determining canonical network equivalence. PLoS One 3(4): e0002051
- Jacobs JA, Frickel S (2009) Interdisciplinarity: A critical assessment. Annu Rev Sociol 35(1):43–65
- Kretschmer H, Rousseau R (2001) Author inflation leads to a breakdown of Lotka's law. J Am Soc Inf Sci Technol 52(8):610–614

- Kolodny O, Feldman MW, Creanza N (2018) Integrative studies of cultural evolution: crossing disciplinary boundaries to produce new insights. Philos Trans R Soc B 373:20170048
- Laland K, Uller T, Feldman M, Sterelny K, Müller GB, Moczek A, Jablonka E, Odling-Smee J, Wray GA, Hoekstra HE, Futuyma DJ, Lenski RE, Mackay TFC, Schluter D, Strassmann JE (2014) Does evolutionary theory need a rethink? Nature 514(7521):161–164
- Landherr A, Friedl B, Heidemann J (2010) A critical review of centrality measures in. Social Netw, Bus Inf Syst Eng 2(6):371–385
- Landry R, Traore N, Godin B (1996) An econometric analysis of the effect of collaboration on academic research productivity. High Educ 32(3):283–301
- Larivière V, Gingras Y, Sugimoto CR, Tsou A (2014) Team size matters: collaboration and scientific impact since 1900. J Assoc Inf Sci Technol 66 (7):1323–1332
- Larivière V, Haustein S, Börner K (2015) Long-distance interdisciplinarity leads to higher scientific impact. PLoS One 10(3):e0122565
- Leahey E, Beckman CM, Stanko TL (2017) Prominent but less productive. Adm Sci Q 62(1):105–139
- Lee S, Bozeman B (2005) The impact of research collaboration on scientific productivity. Soc Stud Sci 35(5):673–702
- Lee Y-N, Walsh JP, Wang J (2015) Creativity in scientific teams: unpacking novelty and impact. Res Policy 44(3):684–697
- Lewens T (2015) Cultural evolution: Conceptual challenges. Oxford University Press: Oxford
- Leydesdorff L, Carley S, Rafols I (2012) Global maps of science based on the new Web-of-Science categories. Scientometrics 94(2):589–593
- Leydesdorff L, Rafols I (2011) Indicators of the interdisciplinarity of journals: diversity centrality, and citations. J Informetr 5(1):87–100
- Liu P, Xia H (2015) Structure and evolution of co-authorship network in an interdisciplinary research field. Scientometrics 103(1):101–134
- Lotka AJ (1926) The frequency distribution of scientific productivity. J Wash Acad Sci 16(12):317?–323
- Mace R (2014) Human behavioral ecology and its evil twin. Behav Ecol 25 (3):443-449
- Mason WA, Jones A, Goldstone RL (2005) Propagation of innovations in networked groups. J Exp Psychol Gen 137(3):422-433
- Mayr E (1982) The growth of biological thought: Diversity, evolution, and inheritance. Belknap Press: Cambridge
- Mesoudi A (2015) Cultural evolution: a review of theory findings and controversies. Evol Biol 43(4):481–497
- Mesoudi A (2017) Pursuing Darwinas curious parallel: Prospects for a science of cultural evolution. Proc Natl Acad Sci 114(30):7853–7860
- Mesoudi A, Whiten A, Laland KN (2006) Towards a unified science of cultural evolution. Behav Brain Sci 29(4):329–383
- National Institutes of Health (2016) NIH-wide strategic plan: Fiscal years 2016–2020. Technical report
- National Science Foundation (2014) Investing in Science, engineering, and education for the nation's future: NSF Strategic Plan for 2014–2018. Technical
- Novak E, Zhao W, Reiser RA (2014) Promoting interdisciplinary research among faculty. J Fac Dev 28(1):19-24
- Perianes-Rodriguez A, Waltman L, van Eck NJ (2016) Constructing bibliometric networks: a comparison between full and fractional counting. J Informetr 10 (4):1178–1195
- Pigliucci M (2006) Phenotypic plasticity and evolution by genetic assimilation. J Exp Biol 209(12):2362–2367
- Piso Z, O'Rourke M, Weathers KC (2016) Out of the fog: catalyzing integrative capacity in interdisciplinary research. Stud Hist Philos Sci Part A 56:84–94
- Porter AL, Rafols I (2009) Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. Scientometrics 81(3):719–745
- Pritchard A (1969) Statistical bibliography or bibliometrics? J Doc 25(4):348–349 Pudovkin AI, Garfield E (2002) Algorithmic procedure for finding semantically related journals. J Am Soc Inf Sci Technol 53(13):1113–1119
- Rafols I, Porter AL, Leydesdorff L (2010) Science overlay maps: a new tool for research policy and library management. J Am Soc Inf Sci Technol 61 (9):1871–1887
- Raina D, Gupta BM (1998) Four aspects of the institutionalization of physics research in India (1990–1950): Substantiating the claims of histortical sociology through bibliometrics. Scientometrics 42(1):17–40
- Rao IKR (1980) The distribution of scientific productivity and social change. J Am Soc Inf Sci 31(2):111–122
- Richerson P, Boyd R (2005) The origin and evolution of cultures. Oxford University Press: New York
- Rousseau R (1992) Breakdown of the robustness property of Lotka's law: the case of adjusted counts for multiauthorship attribution. J Am Soc Inf Sci 43 (10):645–647

- Ruiz-Castillo J, Costas R (2014) The skewness of scientific productivity. J Informetr 8(4):917-934
- Schilling MA, Phelps CC (2007) Interfirm collaboration networks: the impact of large-scale network structure on firm innovation. Manag Sci 53(7):1113–1126 Smith EA, Mulder MB, Hill K (2001) Controversies in the evolutionary social sciences: a guide for the perplexed. Trends Ecol Evol 16(3):128–135
- Stirling A (2007) A general framework for analysing diversity in science technology and society. J R Soc Interface 4(15):707–719
- Sweileh WM, Aİ-Jabi SW, Sawalha AF, Zyoud SH (2016) Bibliometric profile of the global scientific research on autism spectrum disorders SpringerPlus 5 (1):1480
- Sweileh WM, Shraim NY, Al-Jabi SW, Sawalha AF, Rahhal B, Khayyat RA, Zyoud SH (2016) Assessing worldwide research activity on probiotics in pediatrics using Scopus database: 1994–2014. World Allergy Organ J 9(1):25
- Taşkin Z, Aydinoglu AU (2015) Collaborative interdisciplinary astrobiology research: a bibliometric study of the NASA Astrobiology Institute. Scientometrics 103(3):1003–1022
- van Eck NJ, Waltman L (2007) VOS: a new method for visualizing similarities between objects, studies in classification data analysis, and knowledge organization. Springer, Berlin, Heidelberg, p 299–306
- van Eck NJ, Waltman L (2009) Software survey: VOSviewer a computer program for bibliometric mapping. Scientometrics 84(2):523–538
- Wagner CS, Roessner JD, Bobb K, Klein JT, Boyack KW, Keyton J, Rafols I, Börner K (2011) Approaches to understanding and measuring interdisciplinary scientific research (IDR): a review of the literature. J Informetr 5(1):14–26
- Waltman L, van Eck NJ, Noyons EC (2010) A unified approach to mapping and clustering of bibliometric networks. J Informetr 4(4):629-635
- Wang M, Zhao Y, Zhang B (2015) Efficient test and visualization of multi-set intersections. Sci Rep 5(1):16923
- Wimsatt WC (2013) Articulating Babel: an approach to cultural evolution, studies in history and philosophy of science part c: studies in history and philosophy of biological and biomedical. Sciences 44(4):563–571
- Yeung AWK, Goto TK, Leung WK (2017) The changing landscape of neuroscience research 2006–2015: a bibliometric study. Front Neurosci 11:120

Data availability

The data and R code used in the current study, as well as an interactive version of Fig. 2, are available in the Dataverse repository: https://doi.org/10.7910/DVN/LBIDEL.

Acknowledgements

Funding was provided by the John Templeton Foundation Grant ID 60459.

Additional information

Supplementary information: The online version of this article (https://doi.org/10.1057/s41599-018-0175-8) contains supplementary material, which is available to authorized users.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://www.nature.com/reprints

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018