Austral
ECOLOGY A Journal of ecology in the Southern Hemisphere

Austral Ecology (2022) 47, 5–16

What's hot and what's not – Identifying publication trends in insect ecology

NIGEL R. ANDREW,*[1] MALDWYN J. EVANS,[2,3] LAUREN SVEJCAR,[4] KIT PRENDEGAST,[5] LUIS MATA,[6] HELOISE GIBB,[7] MARISA J. STONE[8] AND PHILIP S. BARTON[9]

[1]*Insect Ecology Lab, Zoology, Natural History Museum, University of New England, Armidale, New South Wales, 2351, Australia (Email: nigel.andrew@une.edu.au);* [2]*Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan;* [3]*Fenner School of Environment and Society, The Australian National University, Canberra, Australian Capital Territory, Australia;* [4]*USDA-Agricultural Research Services, Burns, Oregon, USA;* [5]*School of Molecular and Life Sciences, Curtin University, Perth, Bentley, Western Australia, Australia;* [6]*School of Ecosystem and Forest Sciences, The University of Melbourne, Richmond, Victoria, Australia;* [7]*Department of Ecology, Environment and Evolution and Research Centre for Future Landscapes, School of Life Sciences, La Trobe University, Bundoora, Victoria, Australia;* [8]*Environmental Futures Research Institute, School of Environment and Science, Griffith University, Nathan, Queensland, Australia; and* [9]*School of Science, Psychology and Sport, Federation University Australia, Mount Helen, Victoria, Australia*

Abstract Research disciplines in science have historically developed in silos but are increasingly multidisciplinary. Here, we assessed how the insect ecology literature published in ecological and entomological journals has developed over the last 20 years and which topics have crossed discipline boundaries. We used structural topic modelling to assess research trends from 34 304 articles published in six ecology journals and six entomology journals between 2000 and 2020. We then identified and compared topics that emerged from the entire body of literature, or corpus, with topics that emerged from a subsection of articles that focused only on insects (insect corpus). We found that, within the entire corpus, topics on 'Community ecology', 'Traits, life history & physiology' and 'Ecological methods & theory' became more prevalent over time (hot topics), whereas 'Population modelling', 'Insect development', 'Reproduction & ontogeny' and 'Plant growth' declined in prevalence over the 20 years we surveyed (cold topics). In the insect corpus, we found that hot topics included 'Thermal tolerance' and 'Disease vectors', whereas cold topics included 'Herbivore phenology', 'Insect-plant interactions' and 'Parasitoids and parasites'. 'Landscape ecology' was a growth topic area for both corpora. Our findings suggest that insect-related research is a major component of the broader ecological discipline, and there are topics in ecology where insect research aligns with general ecological trends. However, specific topics unique to the insect corpora – such as insect taxonomy – are fundamental to both insect and ecology research. Abstract in Spanish and Portugese is available with online material.

Key words: entomology, insects, publication trends, review, structural topic modelling, taxonomy.

INTRODUCTION

Like all scientific disciplines, ecology changes and evolves (Carmel *et al.* 2013; Borrett *et al.* 2014; Westgate *et al.* 2015). For example, in *Austral Ecology*, several research topics have emerged and changed in importance over the past four decades (Westgate *et al.* 2020): most dramatically, research related to 'communities' and 'landscapes', which were ranked 30 and 29, respectively, in the 1970s (according to the prevalence of keywords among articles) were ranked second and fourth, respectively, in the 2010s. It is intriguing to know where research trends are heading and how foci change as new theories arise, techniques develop and problems emerge. Assessing broad research trends enables researchers to identify gaps and opportunities (e.g. Andrew *et al.* 2013) and to target key research areas (Morton *et al.* 2009; Sutherland *et al.* 2013).

Historically, studies identifying trends and overall scientific literature patterns have relied on classic systematic reviews and meta-analyses (Culina *et al.* 2018; Gurevitch *et al.* 2018). These approaches generally provide an overall assessment of the literature or examine a set of defined questions deemed important by the authors. However, a major shortcoming of systematic literature reviews is that they can be

*Corresponding author.
Accepted for publication April 2021.

doi:10.1111/aec.13052

biased towards the authors' interests or their approach to identifying and selecting relevant literature (Eysenck 1994; Ioannidis 2016; Haddaway *et al.* 2020). Meta-analyses are also prone to conscious and unconscious biases of both researchers and the studies included in analyses and mismatches between data sources (de Vrieze 2018; Reid *et al.* 2018). Here, we wanted to minimise these apparent biases by using a structural topic modelling (STM) approach that uses machine learning text-analysis and provides a more objective assessment of research studies through time (Westgate *et al.* 2015). The STM approach also identifies frequently occurring combinations of words, identified as 'topics' (Westgate *et al.* 2020).

The most significant advantage of using a text-analysis approach is time savings and reduced effort (Marshall & Wallace 2019). Although traditional reviews can address specific questions (e.g. Carmel *et al.* 2013), they can be cumbersome and time-consuming when dealing with broad topics or questions – medical reviews from registration to publication on average take 67 weeks (Borah *et al.* 2017), and environmental systematic reviews take on average 23 weeks (167 days: Haddaway & Westgate 2019). By contrast, a text-analysis approach can quickly assess very large numbers of papers (Marshall & Wallace 2019) and reduce some of the mistakes and biases that can emerge from systematic reviews and meta-analyses (Millard *et al.* 2015). Additionally, text-analysis approaches enable trends to emerge organically, rather than imposing *a priori* restrictions on the data interpretations. Machine learning lets the corpus (i.e. an academic body of literature) do the talking instead of authors 'shoehorning' papers into topics they have decided upon (Westgate *et al.* 2015). This free-form approach provides a more objective overview than traditional review methods such as systematic reviews and meta-analyses (Westgate *et al.* 2015): it allows for unforeseen categories of data to emerge.

One useful text-analysis tool for the ecological sciences is topic modelling (Blei *et al.* 2003; Westgate *et al.* 2015). Topic modelling uses machine learning to identify a given number of topics within a corpus that each contain publications that share similar co-occurrences of words. Conceptually, it is a way of characterising articles' content in a corpus (Murakami *et al.* 2017) according to their similarities and differences. Topics are driven by the data, rather than the researcher (Luiz *et al.* 2019) allowing a more objective categorisation and therefore are a truer reflection of the whole corpus than researcher-determined categorisations. Topic modelling offers a chance for a 'bird's-eye-view' (Luiz *et al.* 2019) of the research field of interest: it allows interrogation of the resultant topics such as their similarities, popularity and growth through time and specificity/generality. It also enables a clearer topic definition, both in terms of thematic content and the articles in which they appear (Westgate *et al.* 2015). This is of interest given that the unification of disparate concepts within scientific fields is considered important in the quest for scientific progress (Chen *et al.* 2009). Here, we assess an ecology corpus and entomology corpus to identify how insect ecology-related research fits into the broader ecological literature and assess changes in key topics over time.

In ecology, classic review approaches have potentially skewed our understanding of key themes and misweighted certain topics/disciplines' importance. For example, taxonomic bias is a key problem in the ecological literature, with 'taxonomic chauvinism' being a critical issue (Bonnet *et al.* 2002; Leather 2009; Troudet *et al.* 2017). These biases in research subjects reflect the researcher's interests and reflect societal interests (Wilson *et al.* 2007; Troudet *et al.* 2017). These biases can also be reflected across a range of specific ecological topics that have been studied across different geographic regions (Culumber *et al.* 2019). As the most diverse eukaryotic group on earth, insects should dominate the ecological literature; however, this not always the case (Troudet *et al.* 2017).

Entomology and insect ecology has had a rich history of developing and pushing ecological research and theory (Leather 2015). Key ecological concepts including mimicry – Batesian (Bates 1981) and Müllerian (Maran 2017); functional responses of organisms (Nicholson 1933); island biogeography (Darlington 1943); succession (Michaud *et al.* 2015); and population ecology (Andrewartha & Birch 1954); all have strong entomological foundations. Understanding recent developments in existing insect ecology research are thus critical to the progression of both entomological and ecological research. Here we were interested in assessing how the influence of entomological research has further developed into the wider ecological sphere.

Culturomics – here defined as publication trends identified through quantitative analysis of word usage in digitised texts (Michel *et al.* 2011) – is a novel and, as yet, not extensively used tool in ecological research. We know of four other reviews that have assessed general ecological research: a traditional review of 750 articles (primarily from abstracts) in eight journals over 30 years (Carmel *et al.* 2013); a machine learning assessment of 84 841 articles (titles, keywords, and abstracts) across 33 high impact factors journals over 40 years (McCallen *et al.* 2019); a machine learning assessment of 32 000 articles from 16 journals between 2000 and 2014 assessing the diverging fields of conservation biology and ecology (Hintzen *et al.* 2020); and a topic modelling assessment of 2 778 full-text articles in a single journal, *Austral Ecology*, across 44 volumes (1976 – 2019) (Westgate *et al.* 2020). However, we know of no other efforts to assess trends across ecological disciplines or crossover of publishing influences and trends.

doi:10.1111/aec.13052

In this study, we investigated if research topics using insects as a focal taxon was similar or different to research found in the wider ecological literature. Our goal was to identify if insect research is biased towards the topics it addresses, and if topics differed between the insect-focused corpus (any article focusing on insects) and the broader ecological corpus (all articles ignoring taxonomic focus). Based on the accessibility of journal abstracts and consistency in format, we restricted the corpus search to the last two decades of research between January 2000 and May 2020. Using text data derived from research abstracts in our search, we addressed the following four questions to address our main goal:

1. What are the key topics found across the entire corpus?
2. What are the dominant taxonomic classes occurring in the key topics of the entire corpus?
3. Are there differences in key topics of study between the ecology and entomology journals in the entire corpus?
4. What are the key topics of research in the insect-specific corpus, and do they differ from topics covered in the entire corpus, or specifically between ecology or entomology journals?

METHODS

We chose 12 journals to form the overall corpus for our study (Appendix S1): six ecology journals and six entomology journals. The journals were: *Ecology*, *Oecologia*, *Oikos*, *Functional Ecology*, *Journal of Animal Ecology*, *Austral Ecology*, *Environmental Ecology*, *Entomologia Experimentalis et Applicata*, *Ecological Entomology*, *Insect Science*, *Austral Entomology/Australian Journal of Entomology*, and *Insect Conservation and Diversity*. These 12 journals were selected as they have a good representation entomological research, are of a high quality and standard and spanned a sufficient number of years for our analysis.

We downloaded the bibliometric data for each journal from the Web of Science in batches of 500 references in June/July 2020. We text-mined abstracts of all articles published between January 2000 and May 2020 in the corpus. For each reference, we extracted: authors; year of publication; journal; volume; title; abstract; keywords. We saved and sequentially named data files to allow batching using R script later in the process. In total, we extracted 34 304 articles which we then classified into the two main groups of journals – ecological = 24 032, and entomological = 10 272; Appendix S1 and S2.

We used the 'taxize' package (Chamberlain & Szöcs 2013) in R (R_Core_Team 2020) to extract the taxonomic names for each article within the whole corpus. Taxize interacts with the Global Names Recognition and Discovery application programming interface (API) to match given text to known taxa. In our case, we used the scrapenames() function to detect all taxonomic terms in all abstracts in the corpus and used the tax_name() function to fetch classes and orders of these terms, using the National Center for Biotechnology Information database API. We then created two corpora – one with all articles (i.e. the entire corpus) and one with only those articles that contained the names of one or more insect taxa in the abstract (i.e. the insect corpus).

To prepare both corpora for analyses, we removed punctuation, numbers and stop words (e.g. and, the, or), from the extracted text of all abstracts and set three characters as the minimum word length. We stemmed words to their root form (e.g. predat = predation, predator, predators, predatory). The final step in data preparation was to remove words that appeared in more than 85% of article abstracts (common words) and words that appeared in fewer than 1% of article abstracts (rare words) – these words provide very little information content (Westgate *et al.* 2015).

Multiple analytical methods for performing topic modelling exist. Two common approaches are the latent Dirichlet allocation (Blei *et al.* 2003) and STM (Roberts *et al.* 2014). Latent Dirichlet allocation (LDA) is perhaps the most common technique for topic modelling (Blei *et al.* 2003). LDA defines topics using sets of words that co-occur at unusual frequencies. Further, articles within the corpus are assumed to consist of multiple topics with weights assigned to each topic for every article (topic weights) (Westgate *et al.* 2015). STM improves on LDA by employing document metadata (e.g. date, journal) to improve the assignment of words to latent topics within a corpus (Roberts *et al.* 2014). This added depth of data allows researchers to make estimations on, for example, the topic relationships to a particular journal type or the trajectory of topics over time (Roberts *et al.* 2016).

We used STM to identify topics within the corpora. STMs allowed us to estimate the relationship between topics and journal type (ecological and entomological) and document metadata such as year of publication (2000-2020) (Roberts *et al.* 2016). This approach enabled us to analyse the trajectory of the topics over the 20-year time frame of the corpora. We fitted STMs by analysing abstracts using the 'stm' package (Roberts *et al.* 2019) in R. When fitting an STM, researchers need to specify the number of topics to identify. We chose 30 topics for both our models, which is the current standard, as this number provided sufficient detail to communicate a summary of the topics in each corpora and contrast the topics with regard to time and journal type without interpretation becoming overly complex (Westgate *et al.* 2015). To quantify the effects of journal type and time, we included a linear 'year' term and a factor 'journal type' term in both of our STMs. We fitted the models using spectral initialisation as it has been found to produce the best results most consistently (Roberts *et al.* 2016).

For both the entire and the insect corpora, we interpreted and labelled our topics (Appendix S3 and S4) by referring to the twenty highest weighted words and also by manually reading abstracts of articles highly associated with each topic. We gave each topic a short title intended to communicate a summary of each topic (Westgate *et al.* 2015). Importantly, the topics identified through our analysis emerge objectively from each corpus, based on the frequency of words in the text among the articles. Similar to an nMDS ordination, we cannot 'choose' or massage these to fit our purposes. A fitted STM outputs a matrix of the $\log_{10}$ probability of occurrence of each analysed word for each topic. We used this matrix to show the

 doi:10.1111/aec.13052

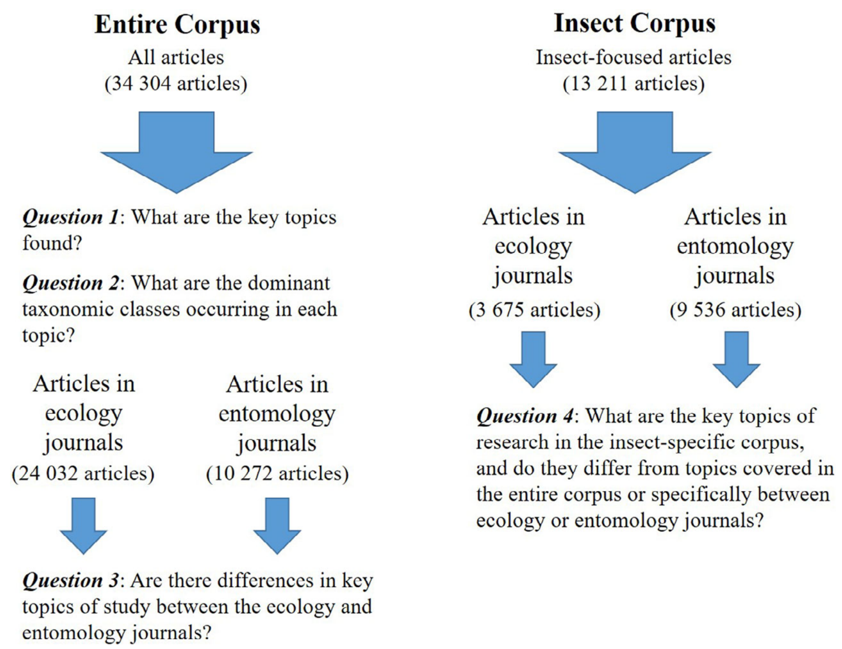

Fig. 1. Overview of study goals and questions.

Euclidean distances between the topics using hierarchical cluster analysis (Westgate *et al.* 2015). Our results show the natural pattern of word co-occurrences in the text and show when they are sufficiently studied to form their own topic. In some instances, text describing particular taxa might occur frequently enough to form their own topic. We did not exclude such taxa as a topic name as it highlights that taxa can indeed be key topics of research in any given field. Additionally, the key bias arises from the cut-off being 30 topics: following Westgate *et al.* (2015).

For each of the four questions, different information was extracted from the dataset (Fig. 1):

Question 1: What are the key topics found across the entire corpus?

We were interested to see how the two journal types were distributed within each of the topics. As such, we fitted topic prevalence between journal type, treating journal type as a factor, using the estimateEffect() function in the ’stm’ package. We were also interested in how topic prevalence changed through time, using this function to fit year as a linear term. We then plotted topics with axes of topic prevalence through time (x-axis) and across the journal type (y-axis).

Question 2: What are the dominant taxonomic classes occurring in the key topics of the entire corpus?

To address question 2, we were interested in identifying the dominant taxa within each topic. We calculated the sum of the topic weights for all articles within the most commonly featured classes (≥ 50 articles) for the entire corpus.

Question 3. Are there differences in key topics of study between the ecology and the entomology journals in the entire corpus?

Here we assessed the 30 most common topics addressed, in both the ecology journals and the entomological journals. We plotted these onto a heatmap ordering topics by their prevalence across the journal types. On the x-axis, ’cold’ topics showed a reduction in prevalence from 2000 to 2020, whereas ’hot’ topics showed an increase. Topics close to the zero line showed no difference over time. On the y-axis, topics close to zero were found equally in both the ecology and entomology journals, and as they moved further from zero, they became more dominant in the specific ’corpus’ analysed. The bars are 95% confidence intervals so we can interpret those topics where the bars cross the zero line as not significant – that is belonging to both equally (Prel *et al.* 2009).

Question 4. What are the key topics of research in the insect-specific corpus, and do they differ from topics covered in the entire corpus or specifically between ecology or entomology journals?

For Question 4, we followed a similar analytical approach to Question 3, but with the subset of articles that specifically focussed on insects (i.e. the insect corpus) wherein we

doi:10.1111/aec.13052

assessed the 30 most common topics addressed in both the ecology journals and the entomological journals. We then plotted the insect corpus results onto heatmaps ordering topics by their prevalence across the journal types.

RESULTS

What are the key topics found across the entire corpus?

We identified 30 clearly defined topics that spanned multiple disciplines of ecology (Fig. 2; Appendix S3; Appendix S5). Eight topics exhibited a positive change over the last 20 years (termed 'hot' topics) – the three with the largest prevalence increase included 'Community ecology', 'Traits, life-history & physiology' and 'Ecological methods & theory'. Eight exhibited no change. Fourteen topics exhibited a reduction (termed 'cold' topics) – in particular 'Population modelling', 'Insect development', 'Reproduction & ontogeny' and 'Plant growth'.

What are the dominant taxonomic classes occurring in the key topics of the entire corpus?

Insect and plant articles dominated the overall corpus that we assessed (Fig. 3). Dominant insect (Class Insecta) topics included 'Insect development', 'Insect genetics' and 'Insect taxonomy & distribution' as well as agricultural pest topics such as 'Insect pest

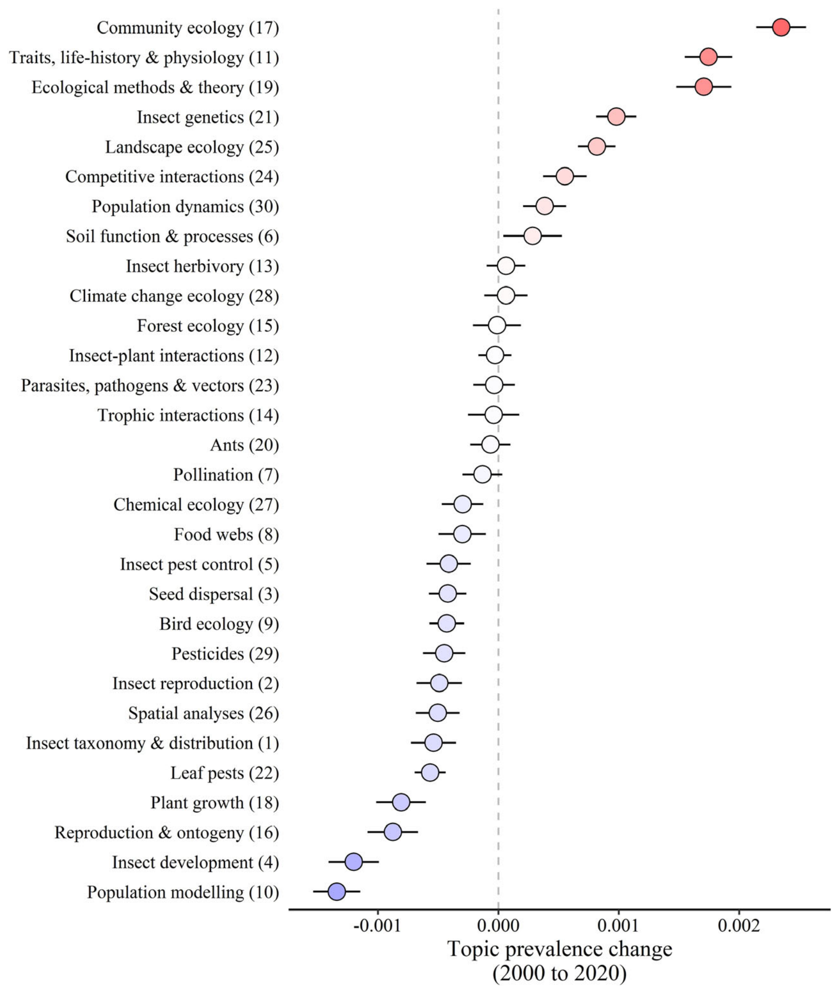

Fig. 2. Change in topic prevalence over time for the whole corpus. Values estimated using the estimate effect() function in the STM model. Each point represents the slope of the prevalence trend over time. Positive values (red, hot) indicate an increase in prevalence over time, and negative values (blue, cold) indicate a decrease in prevalence over time relative to the whole corpus. Error bars represent 95% confidence intervals. Numbers in parenthesis identify the topic number (Appendix S3).

doi:10.1111/aec.13052

control', 'Leaf pests', 'Pesticides' and 'Chemical ecology'. The insect topics were also prevalent in 'Climate change ecology'.

Flowering dicotyledon plants (Class Magnoliopsida) were dominant in topics including 'Leaf pests', 'Insect herbivory', 'Plant growth', 'Soil function' and 'Forest ecology' (Fig. 3). Birds (Class Aves) were most dominant in the topics covering 'Bird ecology', 'Reproduction & ontogeny', 'Population dynamics' and 'Traits, life-history & physiology'. Mammals (Class Mammalia) were dominant in the topics covering 'Reproduction & ontogeny', 'Population dynamics', 'Trophic interactions', 'Ecological methods & theory', 'Spatial analysis' and 'Population modelling'.

Are there differences in key topics of study between the ecology and the entomology journals?

Fifteen topics were more prevalent in ecology journals, 11 topics more prevalent in the entomology journals, and four topics overlapped between the two (Appendix S5 and Appendix S6). The three 'hot topics' (i.e. strongest growth areas) in the Ecology corpus were 'Community ecology', 'Traits, life-history & physiology' and 'Ecological methods & theory' (Fig. 4). For the Entomological journals, the hottest topic was 'Insect genetics'. 'Landscape ecology' was a hot topic across both ecology and entomology journals with 'Insect herbivory', 'Myrmecology' and 'Pollination' staying relatively static in topic prevalence over time. 'Population modelling' was the coldest topic in the ecology journals, whilst 'Insect development' was the coldest topic in the entomology journals.

What are the key topics of research in the insect-specific corpus and do they differ from topics covered in the entire corpus or specifically between ecology or entomology journals?

We identified 30 clearly defined topics that spanned multiple insect corpus topics (Appendix S4), with ten topics increasing in prevalence, 13 staying relatively static, and seven reducing in prevalence from 2000 to 2020 (Appendix S7). Ten of these topics were predominantly found in ecology journals (Appendix S8). Insect-only hot topics in the ecology journals included 'Trophic interactions', 'Community ecology', 'Life-history traits' and 'Herbivory responses'. Cold topics included 'Population dynamics' and 'Predation' (Fig. 5). Fourteen topics were predominantly in entomology journals (Appendix S4

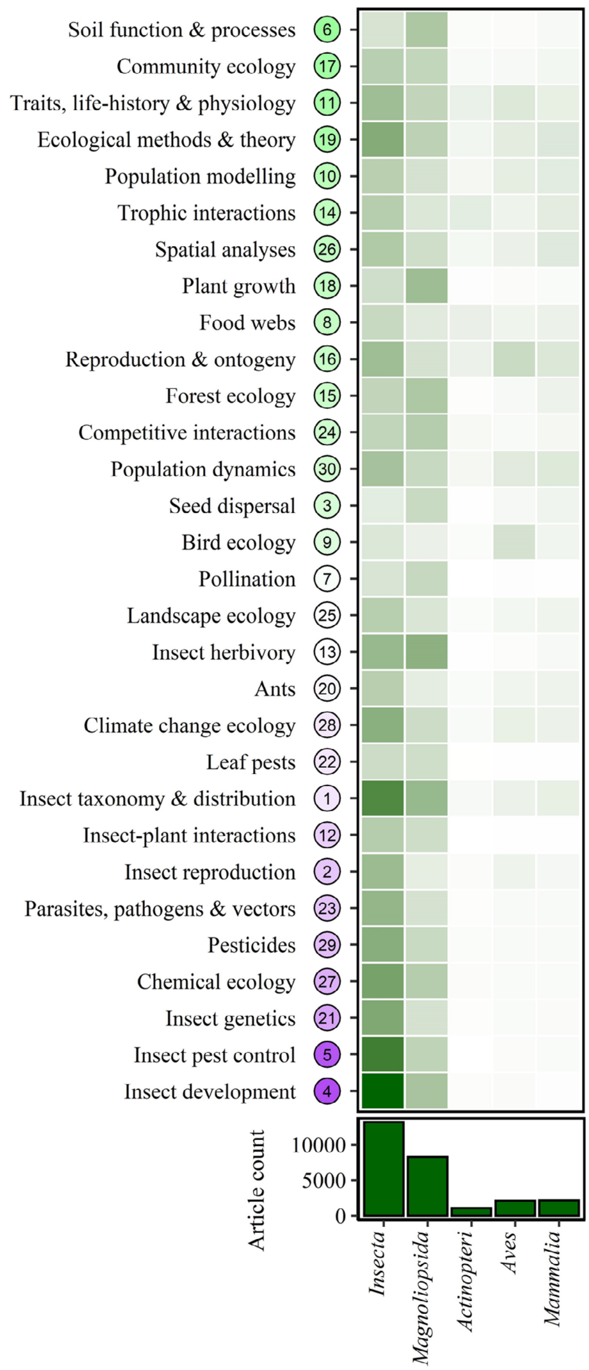

Fig. 3. Heatmap of classes across the corpus. Includes classes that appeared in 1000 or more articles. Classes are clustered according to topic similarity. Topics are ordered according to their journal type prevalence (green coloured circles representing topics prevalent with ecological journals and purple circles indicating prevalence with entomological journals; see Appendix S5), and the number within the in circle identifies the topic number. Cell shade represents the sum of the topic weights for all articles that contain that class. Point colours along the y-axis represent the journal type prevalence of the topics.

doi:10.1111/aec.13052

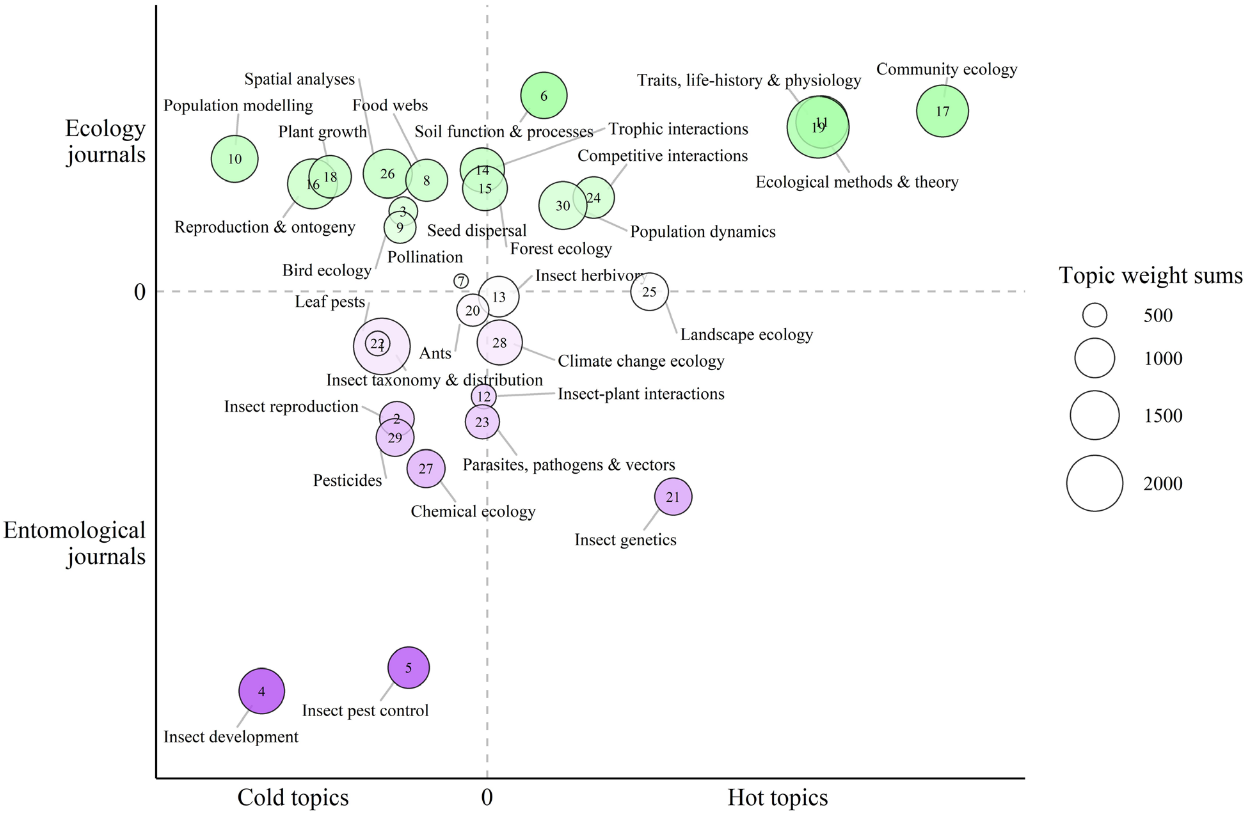

Fig. 4. Plot of topics according to prevalence over time (x axis) and journal type (y axis). The number within the circle identifies the topic number, and size of the circle represents the sum of the article weights for each topic (Appendix S3).

and Appendix S9). The hottest topic in the entomology journals was clearly 'Gene expression', with smaller increases in 'Crop pests in soybean, potato & cotton' and 'IPM-Biological control'. Entomological journal's cold topics included 'Parasitoids & parasites', 'Phenology & development' and 'Noctuidae pests' (Fig. 4). Three topics were common to both journal groups (Appendix S8). The primary hot topic across all journals (Fig. 5) was 'Thermal tolerance'. 'Nutrition' had a similar influence over time; while 'Insect-plant interactions' was a 'cold' topic (Fig. 5).

Across the journal type and topics, five insect orders dominated the corpora (Fig. 6): Lepidoptera, Hymenoptera, Hemiptera, Coleoptera and Diptera. All five orders made strong contributions to most of the 30 topics identified.

DISCUSSION

Our study accomplished two main goals: 1) we identified how insect ecology-related research fits into the broader ecological literature and 2) we assessed how changes in key topics over time differ between literature in ecological and entomological journals. We found that three topics from the entire corpus increased in their prevalence over the last 20 years: 'Community ecology', 'Traits, life-history & physiology' and 'Ecological methods & theory'. Previous text analysis research published in *Austral Ecology* also found 'Community ecology' was a dominant topic over the last decade (Westgate *et al.* 2020), while a review of 33 high impact journals found 'Traits' to be a similarly prevalent topic (McCallen et al 2019). However, McCallen *et al.* (2019) found 'Climate change' and 'Genetics' to be the other two major topics in their analyses. The difference between their findings and ours could be the result of a heavy bias for high-impact global studies, such as climate change and medical research.

Key topics from the entire corpus that declined in their prevalence over time included 'Population modelling', 'Insect development', 'Reproduction & ontogeny' and 'Plant growth'. These topics may be picked up by other journals not assessed here, or their influence, or terminology used to refer to these topics, may have evolved. For example, population modelling may have morphed into other types of modelling, as the topic word 'model' was also included in both the topics' Ecological methods &

 doi:10.1111/aec.13052

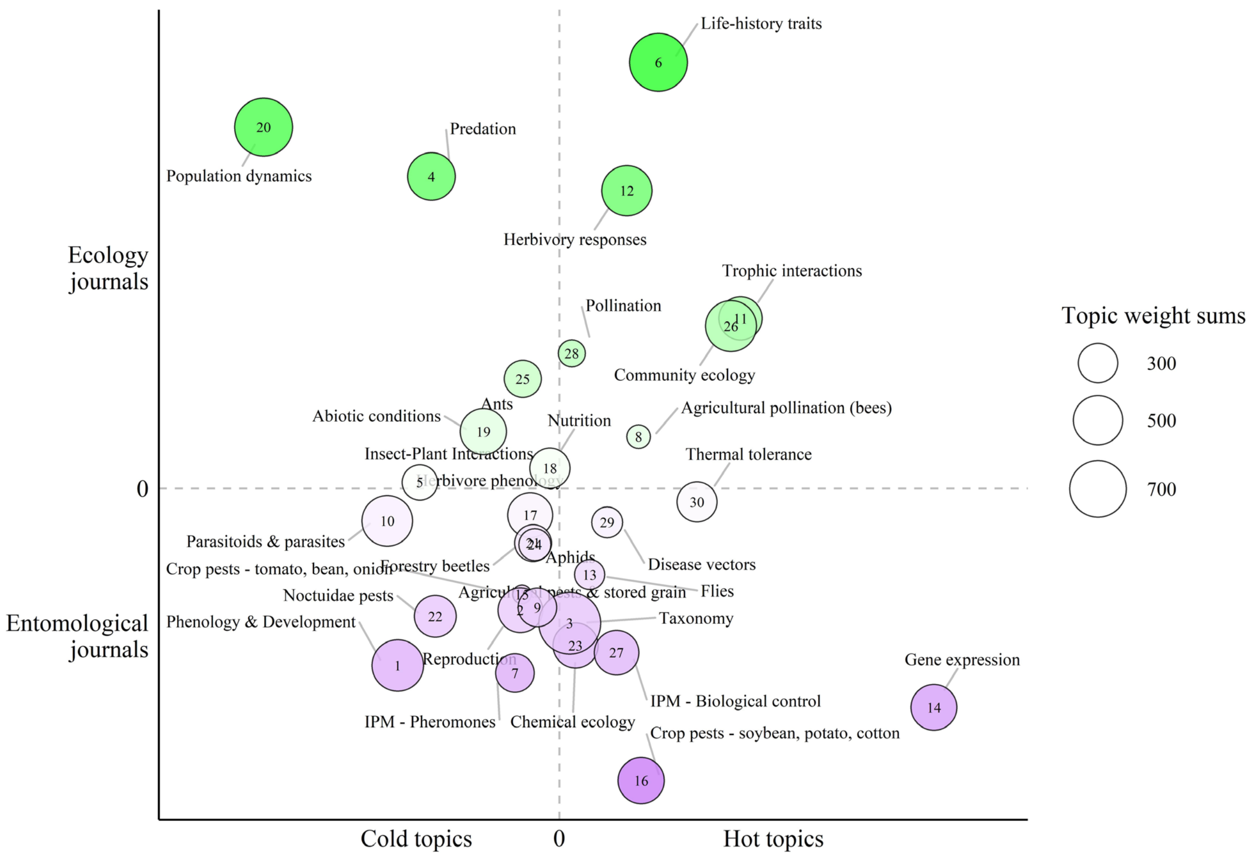

Fig. 5. Plot of topics according to prevalence over time (x-axis) and journal type (y-axis) for the insect-only corpus. The number within the circle identifies the topic number, and size of the circle represents the sum of the article weights for each topic (Appendix S4).

theory' (Topic 19) and 'Spatial analysis' (Topic 26) (Appendix S3). However, this could be a wider trend as the themes' Life-history', 'Plant reproduction', 'Survivorship', 'Food webs', 'Carrying capacity', 'Plant physiology', 'Seasonal trends', 'Development' and 'Herbivory' all reduced in prevalence across 33 high-impact journals assessed by McCallen *et al.* (2019). Additionally, there has been an increase in the number of journals in both ecology (number of ecology journals in 2019 – 169; 2000 – 100; (Clarivate_Analytics 2020)) and entomology (number of entomology journals in 2019 101; 2000 – 66; (Clarivate_Analytics 2020)) disciplines that have been associated with an exponential increase in publications broadly in the natural sciences (Bornmann & Mutz 2015). In other broad discipline areas, such as higher education research, this has led to splitting of research into isolated 'islands' and further discipline-specific divisions because of specialisation (Daenekindt & Huisman 2020).

Insects and plants were the dominant taxa assessed in this study (Class Insecta: 33.6% of papers; Class Magnoliopsida 21.2%), followed by Class Aves (6.1%) and Class Mammalia (5.8%). As we targeted insects, this is to be expected. However, the bias toward insects in publications is unusual with a general bias towards endotherms for both biodiversity occurrence data (Troudet *et al.* 2017), ecological research (Leather 2009) and behavioural research (Rosenthal *et al.* 2017).

Distinct topics from the entire corpus tended to be more prevalent in either ecology or entomology journals, but some topics were prevalent in both. For example, 'Landscape Ecology' was a growth topic area for both journal categories, suggesting that larger spatial scale research is increasingly studied. Insect–plant interaction research, specifically 'Insect herbivory' and 'Pollination', were both prevalent over the last 20 years in both journal categories. This crossover may be because of how the two topics explicitly involve insects interacting with other non-insect taxa, specifically with plants. Additionally, research on 'Ants' is a dominant topic published equally in ecology and entomology journals through time. This suggests that not only is the larger-scale ecological literature influencing the entomological

doi:10.1111/aec.13052

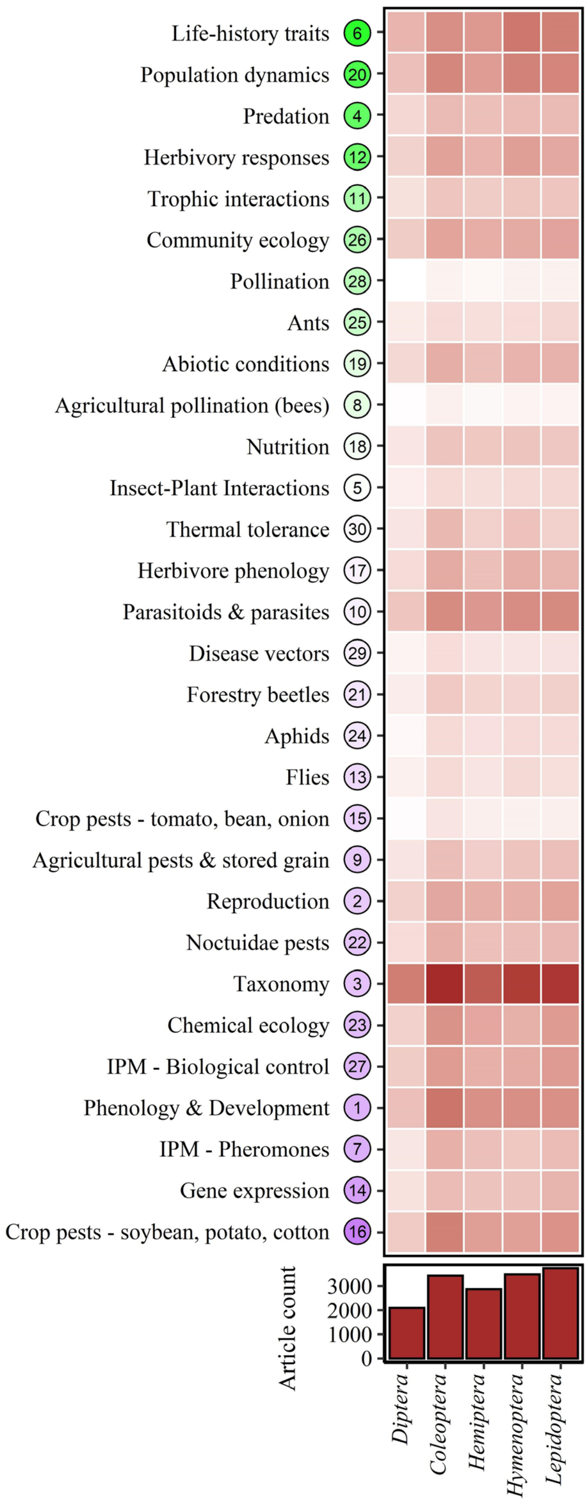

Fig. 6. Heatmap of orders across the insect-only corpus. Includes orders that appeared in 1000 or more articles. Orders are clustered according to topic similarity. Topics are ordered according to their journal type prevalence (green coloured circles representing topics prevalent with ecological journals and purple circles indicating prevalence with entomological journals; see Appendix S7), and number within the circle identifies the topic number. Cell shade represents the sum of the topic weights for all articles that contain that class. Point colours along the y-axis represent the journal type prevalence of the topics (Appendix S4).

literature, but there is also insect research reaching into the broader ecological literature.

Within the insect-specific corpus, hot topics crossing both entomology and ecology journals include ’Thermal tolerance’ and ’Disease vectors’. Topics with similar influence over time include ’Forestry beetles’ and ’Nutrition’. The ’Thermal tolerance’ topic broadly covers climate change research, range shift, latitudinal gradients and cold resistance (Appendix S4). Researchers are moving past the generic climate change topics and becoming more nuanced in assessing the direct impacts of human-induced climate change across various disciplines (Haunschild *et al.* 2016; Wang *et al.* 2018; Callaghan *et al.* 2020). Work on critical thermal limits – both upper and lower – are key areas of this work. Within the climate change framework, it is also surprising that ’Population dynamics’ research has reduced in its emphasis in the ecology journals. We anticipated these critical aspects of life history would continue to be well studied. Gene expression is the key growth area in the entomology journals. This increase is likely because of gene editing, sequencing and DNA identification becoming more accessible in terms of both cost and availability (e.g. Cruaud *et al.* 2019; Klein & Hainer 2020; Leung *et al.* 2020).

Overall, we found positive evidence indicating insect-specific research is overlapping with the broader ecological research areas and continuing to be an active research topic, or increasing in prevalence, over time. In particular, ’Landscape ecology’, ’Insect herbivory’ and ’Ant’ research were strong and active throughout the major ecological and entomological journals we assessed. There was also a range of topics unique to each corpus. For the insect corpus ‘Traits’, ‘Interactions’ and community-level research increased in ecological journals, whereas more genetic and pest-based analysis were key growth areas in the entomological journals. Even though we only assessed a snapshot of journals, we would expect these trends to follow if other ecological and entomological journals were to be included in a broader assessment, as they have been identified in previous ecological literature analyses (McCallen *et al.* 2019).

We expected to see a decrease in entomological taxonomy research over time because of decreases in paid taxonomic job positions in Australia (Braby & Williams 2016; Weaver 2017). However, we found a dominance and slight increase in taxonomic research

 doi:10.1111/aec.13052

in entomology journals over the past 20 years with taxonomy considered a potentially 'warm' topic. Interestingly, this increase is independent from changes in genetic research (based on the top 20 topic words, Appendix S4). There have been many reviews and commentaries identifying issues with taxonomic research over the past few decades: from taxonomic bias – both across classes (Troudet *et al.* 2017) and within the Insecta class (Zuk *et al.* 2014), through to the lack of taxonomic skills and the loss of taxonomic expertise, such as for flowering plants (Bebber *et al.* 2014). However, other assessments found increases in taxonomy across all animal taxa (Vinarski 2020). Here we also found that topics related to insect taxonomic research are active and have increased in prevalence over the past 20 years – for the top five insect orders, taxonomy was highly weighted. Such integration between taxonomy and ecology needs continual nurturing and strengthening (Halme *et al.* 2015), enabling opportunities for critical research in new and emerging issues to be acted upon into the future.

CONCLUSION

Our study demonstrates entomological research was and continues to play a major role in the development of ecological concepts and theories. The publication assessment carried out here – although not exhaustive – has extracted some key elements of how topics move between taxa specific (entomology) and broader discipline-based (ecology) journals. Crossover of key ideas and theories from taxa to discipline is important for developing a holistic understanding of the interactions of all components in an ecosystem or ecological function. A holistic understanding of ecological function enables us to move the ecology discipline forward and develop tools to assess and manage 'wicked problems' (Lönngren & Svanström 2016; DeFries & Nagendra 2017; Wohlgezogen *et al.* 2020) that are continually emerging.

ACKNOWLEDGEMENTS

We thank the Ecological Society of Australia for providing the funding and resource to enable the workshop to develop the manuscripts initial ideas. The Ecological Society of Australia, Insect Ecology Research Chapter runs a series of workshops to bring insect ecologists together throughout the year. A key driver of these workshops is articulating the need for ecological research to mobilise at a national scale, as well as integrate between other taxa and ecological disciplines to develop a broader ecological perspective. This study's initial impetus was an insect ecology-focussed workshop on the 30th November 2018, as part of the Ecological Society of Australia annual conference in Brisbane, Australia. Workshop participants identified the critical lack of understanding of how insect-related literature fits into the broader ecological landscape and how resolving this knowledge gap might help identify future insect ecology research priorities. We thank the participants of the workshop who assisted with ideas and impetus for the work: Brogan Amos, Barbara Downes, Kathy Ebert, Edward Fensom, Adam Frew, Lachlan Jones, Thomas Hayashi, Andrew Hayes, Dieter Hochuli, Melissa Houghton, Manual Lequerica, Grace Leung, Andrew Maynard, Jaye Newman, Hazel Parry, Emma Razeng, Thomas Sayers, Manu Saunders, Jessa Thurman, Andrew Wallace-Mitchell, Blythe Vogel and Alan York. We also thank Lucy Johanson for assisting in extracting some records from the Web of Science database. We also thank Elsie Denton and Kirk Davies for reviewing a previous version of the manuscript.

AUTHOR CONTRIBUTION

Nigel Andrew: Conceptualization (equal); Data curation (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (lead); Resources (equal); Writing-original draft (lead); Writing-review & editing (lead). **Maldwyn J Evans:** Formal analysis (lead); Methodology (equal); Software (lead); Visualization (lead); Writing-original draft (equal); Writing-review & editing (equal). **Lauren Svejcar:** Conceptualization (equal); Investigation (equal); Writing-review & editing (equal). **Kit Prendergast:** Conceptualization (equal); Investigation (equal); Writing-review & editing (equal). **Luis Mata:** Conceptualization (equal); Data curation (equal); Investigation (equal); Writing-review & editing (equal). **Heloise Gibb:** Conceptualization (equal); Funding acquisition (equal); Writing-review & editing (equal). **Marisa Stone:** Conceptualization (equal); Data curation (equal); Investigation (equal); Writing-review & editing (equal). **Philip Barton:** Formal analysis (equal); Investigation (equal); Methodology (equal); Writing-original draft (equal); Writing-review & editing (equal).

CONFLICT OF INTEREST

N. R. Andrew is the Editor-in-Chief and P. S. Barton is an Editor of *Austral Ecology*. They played no part in the editorial decisions made relating to this manuscript.

REFERENCES

Andrew N. R., Hill S. J., Binns M. *et al.* (2013) Assessing insect responses to climate change: What are we testing for? Where should we be heading? *PeerJ* **1,** e11.

Andrewartha H. G. & Birch L. C. (1954) *The Distribution and Abundance of Animals*. University of Chicago Press, Chicago.

Bates H. W. (1981) Contributions to an insect fauna of the Amazon valley (Lepidoptera: Heliconidae). *Biol. J. Linn. Soc.* **16,** 41–54.

Bebber D. P., Wood J. R. I., Barker C. & Scotland R. W. (2014) Author inflation masks global capacity for species discovery in flowering plants. *New Phytol.* **201,** 700–6.

Blei D., Ng A. & Jordan M. (2003) Latent dirichlet allocation. *J. Machine Learn. Res.* **3,** 993–1022.

Bonnet X., Shine R. & Lourdais O. (2002) Taxonomic chauvinism. *TREE* **17,** 1–3.

Borah R., Brown A. W., Capers P. L. & Kaiser K. A. (2017) Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open* **7,** e012545.

Bornmann L. & Mutz R. (2015) Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *J. Ass. Inform. Sci. Technol.* **66,** 2215–22.

Borrett S. R., Moody J. & Edelmann A. (2014) The rise of Network Ecology: Maps of the topic diversity and scientific collaboration. *Ecol. Model.* **293,** 111–27.

Braby M. F. & Williams M. R. (2016) Biosystematics and conservation biology: critical scientific disciplines for the management of insect biological diversity. *Austral Entomol.* **55,** 1–17.

Callaghan M. W., Minx J. C. & Forster P. M. (2020) A topography of climate change research. *Nature CC* **10,** 118–23.

Carmel Y., Kent R., Bar-Massada A. *et al.* (2013) Trends in ecological research during the last three decades – A systematic review. *PLoS One* **8,** e59813.

Chamberlain S. & Szöcs E. (2013) taxize: taxonomic search and retrieval in R. *F1000Research* **2,** 191.

Chen C., Chen Y., Horowitz M., Hou H., Liu Z. & Pellegrino D. (2009) Towards an explanatory and computational theory of scientific discovery. *J. Inform.* **3,** 191–209.

Clarivate_Analytics (2020) InCites Journal Citation Reports dataset updated Oct 20, 2020.

Cruaud A., Nidelet S., Arnal P. *et al.* (2019) Optimized DNA extraction and library preparation for minute arthropods: Application to target enrichment in chalcid wasps used for biocontrol. *Mol. Ecol. Resour.* **19,** 702–10.

Culina A., Crowther T. W., Ramakers J. J. C., Gienapp P. & Visser M. E. (2018) How to do meta-analysis of open datasets. *Nat. Ecol. Evol.* **2,** 1053–6.

Culumber Z. W., Anaya-Rojas J. M., Booker W. W. *et al.* (2019) Widespread biases in ecological and evolutionary studies. *Bioscience* **69,** 631–40.

Daenekindt S. & Huisman J. (2020) Mapping the scattered field of research on higher education. A correlated topic model of 17,000 articles, 1991–2018. *High. Educ.* **80,** 571–87.

Darlington P. J. (1943) Carabidae of mountains and islands: Data on the evolution of isolated faunas, and on atrophy of wings. *Ecol. Monogr.* **13,** 37–61.

de Vrieze J. (2018) The metawars. *Science* **361,** 1184–8.

DeFries R. & Nagendra H. (2017) Ecosystem management as a wicked problem. *Science* **356,** 265–70.

Eysenck H. J. (1994) Systematic reviews: Meta-analysis and its problems. *BMJ* **309,** 789–92.

Gurevitch J., Koricheva J., Nakagawa S. & Stewart G. (2018) Meta-analysis and the science of research synthesis. *Nature* **555,** 175–82.

Haddaway N. R., Bethel A., Dicks L. V. *et al.* (2020) Eight problems with literature reviews and how to fix them. *Nat. Ecol. Evol* **4,** 1582–1589.

Haddaway N. R. & Westgate M. J. (2019) Predicting the time needed for environmental systematic reviews and systematic maps. *Conserv. Biol.* **33,** 434–43.

Halme P., Kuusela S. & Juslén A. (2015) Why taxonomists and ecologists are not, but should be, carpooling? *Biodiv. Conserv.* **24,** 1831–6.

Haunschild R., Bornmann L. & Marx W. (2016) Climate change research in view of bibliometrics. *PLoS One* **11,** e0160393.

Hintzen R. E., Papadopoulou M., Mounce R. *et al.* (2020) Relationship between conservation biology and ecology shown through machine reading of 32,000 articles. *Conserv. Biol.* **34,** 721–32.

Ioannidis J. P. A. (2016) The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Quart.* **94,** 485–514.

Klein D. C. & Hainer S. J. (2020) Genomic methods in profiling DNA accessibility and factor localization. *Chromosome Res.* **28,** 69–85.

Leather S. (2009) Taxonomic chauvinism threatens the future of entomology. *Biologist* **56,** 10–3.

Leather S. R. (2015) Influential entomology: a short review of the scientific, societal, economic and educational services provided by entomology. *Ecol. Entomol.* **40,** 36–44.

Leung K., Ras E., Ferguson K. B. *et al.* (2020) Next-generation biological control: the need for integrating genetics and genomics. *Biolog. Rev.* **95,** 1838–54.

Lönngren J. & Svanström M. (2016) Systems thinking for dealing with wicked sustainability problems: Beyond functionalist approaches. In: *New Developments in Engineering Education for Sustainable Development* (eds W. Leal Filho & S. Nesbit) pp. 151–60.Springer International Publishing, Cham.

Luiz O. J., Olden J. D., Kennard M. J. *et al.* (2019) Trait-based ecology of fishes: A quantitative assessment of literature trends and knowledge gaps using topic modelling. *Fish Fish.* **20,** 1100–10.

Maran T. (2017) The structure of mimicry. In: *Mimicry and Meaning: Structure and Semiotics of Biological Mimicry*, Biosemiotics, vol. **18,** pp. 15-34. Springer International Publishing, Cham.

Marshall I. J. & Wallace B. C. (2019) Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Sys. Rev.* **8,** 163.

McCallen E., Knott J., Nunez-Mir G., Taylor B., Jo I. & Fei S. (2019) Trends in ecology: shifts in ecological research themes over the past four decades. *Front. Ecol. Environ.* **17,** 109–16.

Michaud J.-P., Schoenly K. G., & Moreau G. & Handling Editor Daniel D. (2015) Rewriting ecological succession history: did carrion ecologists get there first? *Quart. Rev. Biol.* **90,** 45–66.

Michel J.-B., Shen Y. K., Aiden A. P. *et al.* (2011) Quantitative analysis of culture using millions of digitized books. *Science* **331,** 176–82.

Millard L. A., Flach P. A. & Higgins J. P. (2015) Machine learning to assist risk-of-bias assessments in systematic reviews. *Int. J. Epidemiol.* **45,** 266–77.

 doi:10.1111/aec.13052

Morton S. R., Hoegh-Guldberg O., Lindenmayer D. B. *et al.* (2009) The big ecological questions inhibiting effective environmental management in Australia. *Aust. Ecol.* **34,** 1–9.
Murakami A., Thompson P., Hunston S. & Vajn D. (2017) 'What is this corpus about?': Using topic modelling to explore a specialised corpus. *Corpora* **12,** 243–77.
Nicholson A. J. (1933) The balance of animal populations. *J. Anim. Ecol.* **2,** 132–78.
Prel J.-B., Hommel G., Röhrig B. & Blettner M. (2009) Confidence interval or p-value? Part 4 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt international* **106,** 335–9.
R_Core_Team (2020) *R: A language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Reid J. L., Fagan M. E. & Zahawi R. A. (2018) Positive site selection bias in meta-analyses comparing natural regeneration to active forest restoration. *Sci. Adv.* **4,** eaas9143.
Roberts M. E., Stewart B. M., Tingley D. *et al.* (2014) Structural topic models for open-ended survey responses. *Am. J. Pol. Sci.* **58,** 1064–82.
Roberts M. E., Stewart B. M. & Tingley D. (2016) Navigating the local modes of big data: The case of topic models. In: *Computational Social Science: Discovery and Prediction* (ed R. M. Alvarez) pp. 51–97.Cambridge University Press, Cambridge.
Roberts M. E., Stewart B. M. & Tingley D. (2019) stm: An R Package for Structural Topic Models. *J. Stat. Softw.* **91,** 40.
Rosenthal M. F., Gertler M., Hamilton A. D., Prasad S. & Andrade M. C. B. (2017) Taxonomic bias in animal behaviour publications. *Anim. Behav.* **127,** 83–9.
Sutherland W. J., Freckleton R. P., Godfray H. C. J. *et al.* (2013) Identification of 100 fundamental ecological questions. *J. Ecol.* **101,** 58–67.
Troudet J., Grandcolas P., Blin A., Vignes-Lebbe R. & Legendre F. (2017) Taxonomic bias in biodiversity data and societal preferences. *Sci. Rep.* **7,** 9132.
Vinarski M. V. (2020) Roots of the taxonomic impediment: Is the "integrativeness" a remedy? *Integr. Zool.* **15,** 2–15.
Wang Z., Zhao Y. & Wang B. (2018) A bibliometric analysis of climate change adaptation based on massive research literature data. *J. Clean. Prod.* **199,** 1072–82.
Weaver H. J. (2017) *State of the Science of Taxonomy in Australia: Results of the 2016 Survey of Taxonomic Capacity*. Commonwealth of Australia Canberra, Canberra.
Westgate M., Barton P. S., Lindenmayer D. B. & Andrew N. R. (2020) Quantifying shifts in topic popularity over 44 years of Austral Ecology. *Aust. Ecol.* **45,** 663–71.
Westgate M. J., Barton P. S., Pierson J. C. & Lindenmayer D. B. (2015) Text analysis tools for identification of emerging topics and research gaps in conservation science. *Conserv. Biol.* **29,** 1606–14.
Wilson J. R., Procheş Ş., Braschler B., Dixon E. S. & Richardson D. M. (2007) The (bio)diversity of science reflects the interests of society. *Front. Ecol. Environ.* **5,** 409–14.
Wohlgezogen F., McCabe A., Osegowitsch T. & Mol J. (2020) The wicked problem of climate change and interdisciplinary research: Tracking management scholarship's contribution. *J. Manag. Org.* **26,** 1048–72.
Zuk M., Garcia-Gonzalez F., Herberstein M. E. & Simmons L. W. (2014) Model systems, taxonomic bias, and sexual selection: beyond *Drosophila*. *Ann. Rev. Entomol.* **59,** 321–38.

SUPPORTING INFORMATION

Additional supporting information may/can be found online in the supporting information tab for this article.

Appendix S1. The number of articles in the whole corpus for each journal.

Appendix S2. The number of articles in the whole corpus for each year and journal category.

Appendix S3. Uncovered topics from 34,304 research articles across the whole corpus.

Appendix S4. Uncovered topics from 10,272 research articles across the insect-only corpus.

Appendix S5. Topic prevalence in each of the journal types (Entomological journals v Ecology journals).

Appendix S6. Hierarchical cluster analysis of topic representing topic distances and similarities.

Appendix S7. Topic prevalence over time for the insect-only corpus.

Appendix S8. Topic prevalence in each of the journal types (Entomological journals v Ecology journals) for the insect-only corpus.

Appendix S9. Hierarchical cluster analysis of topic representing topic distances and similarities for the insect-only corpus.

doi:10.1111/aec.13052