CITATION:	E. L, Zvereva., and Kozlov, M. V.. 2019. Biases in studies of spatial patterns in insect herbivory. *Ecological Monographs* 89(3):e01361.DOI:10.1002/ecm.1361 which has been published in final form at
DOI	https://doi.org/10.1002/ecm.1361

Biases in studies of spatial patterns in insect herbivory

Elena L. Zvereva[1,2] and Mikhail V. Kozlov[1]

[1] *Department of Biology, University of Turku, 20014 Turku, Finland*

[2] E-mail: elezve@utu.fi

Abstract The properties of the human mind are responsible for a number of biases that affect the quality of scientific research. However, scientists working in the fields of ecology and environmental science rarely take these biases into account. We conducted a meta-analysis of data extracted from 125 publications comparing woody plant damage by defoliating insects in different environments in order to understand the extent to which our knowledge on spatial patterns in herbivory is affected by various biases. We asked which research methods are most prone to biases and whether these biases lead to overestimation of the effects under study. The effects sizes (ESs) decreased with increases in the numbers of plant species involved in the study, with 61% lower ESs for herbivory estimated on all plants growing in study plots compared to herbivory on selected species. ESs also depended on the leaf sampling procedure: when all leaves from a tree or branch were sampled for measurements of herbivory or when random or systematic selection protocols were applied, ESs were 74% smaller than in cases of more subjective haphazard selection. In addition, ESs were 97 and 135% greater when the person conducting sampling and measuring was aware of the research hypothesis or sample origin, when compared with situations when the observer was blinded to these factors. The impacts of cognitive biases on the study outcomes significantly decreased with the increase in publication year; however, this pattern emerged mostly due to high-ranked journals and was non-significant for other journals. Using the studies of spatial patterns in herbivory as an example, we showed that our ecological and environmental knowledge is considerably biased due to an unconscious tendency of researchers to find support for their hypotheses and expectations, which generally leads to overestimation of the effects under study. Cognitive biases can be avoided by using different methods, such as applying randomization procedures in sampling and blinding of research hypotheses and sample origins. These measures should be seen as obligatory; otherwise, accumulation of the biased results in primary studies may ultimately lead to false general conclusions in subsequent research synthesis.

Key words: cognitive biases, meta-analysis, blind protocols, randomization, research methods, insect herbivory, defoliators, haphazard selection, scientific methodology, spatial patterns in herbivory, temporal trends in research, effect size.

INTRODUCTION

The human mind is undoubtedly the most important tool in scientific research, as only the human mind can interpret the data in order to arrive at a logical conclusion (Leedy and Ormrod 2001). Nevertheless, the properties of the human mind can affect the quality of the research through insertion of a number of biases at the planning, data collection, analysis, and/or publication phases of scientific research (Pannucci and Wilkins 2010). In science, bias is defined as systematic errors in results or inferences that favor one outcome over others (Gluud 2006). Unfortunately, some biases are rarely considered by scientists working in the fields of ecology and evolution.

Publication bias has a recognized influence on the understanding of ecological and evolutionary processes and is widely appreciated (Fanelli 2010, 2012, Jennions et al. 2013). By contrast, the occurrence and importance of biases introduced at pre-publication stages of ecological research have received little attention (but see Kozlov and Zvereva 2009, 2015, Koricheva et al. 2013a, Kozlov et al. 2014). This is especially true for confirmation bias: a tendency to search for, interpret, and favor information in a way that confirms one’s pre-existing beliefs or hypotheses (Forstmeier et al. 2017). Confirmation bias is a well-documented phenomenon in psychology and cognitive sciences; it results primarily from automatic processes that occur unintentionally (Hergovich et al. 2010). Within biological disciplines, confirmation bias has received sufficient attention only in studies of animal behavior (Marsh and Hanlon 2007, van Wilgenburg and Elgar 2013, Traniello and Bakker 2015, Tuyttens et al. 2014, 2016).

Confirmation bias is usually demonstrated by comparing the results of studies conducted blindly (i.e., when the observer was not aware of the research hypothesis being tested) with the results of non-blind studies (i.e., when the observer knew what results might be expected) (van Wilgenburg and Elgar 2013, Holman et al. 2015). Blinding is a routine procedure in medical research, because non-blind protocols were repeatedly found to overestimate the effects of a specific treatment (Noseworthy et al. 1994, Saltaji et al. 2018). At the same time, blinding is only rarely reported in ecological and evolutionary studies (Forstmeier et al. 2017), and the outcomes of blind and non-blind methods have rarely been compared in ecological research. However, when this type of comparison has been conducted, the lack of blinding with respect to the hypothesis being tested or the treatment condition of a sample usually resulted in overestimation of the effects under study by the observers (Kozlov et al. 2014, Kozlov and Zvereva 2015). Consequently, meta-analysis of experimental studies within the life sciences indicates that non-blind studies tend to report higher effect sizes (ESs, hereafter) and more significant P values when compared to blind studies (Holman et al. 2015). We expect that the impacts of various biases will be stronger on research areas that rely on observational data than on experimental data, because collection of observational data may be more prone to unconscious biases. Studies of spatial patterns in herbivory and of environmental factors driving these patterns provide a representative example of this type of research field.

The relationships between plants and herbivores are among the most intensively studied biotic interactions (Tylianakis et al. 2008, Jamieson et al. 2012), and the amount of plant biomass consumed by herbivores is the key characteristic of the intensity of these interactions. Explicitly or implicitly, the numerical values reflecting the pressure imposed by herbivorous insects on plants are among the cornerstones of numerous hypotheses/theories related to insect-plant relationships, such as the ‘green world’ hypothesis (Hairston et al. 1960, Polis 1999), the exploitation ecosystem hypothesis (Oksanen et al. 1981, Polis 1999), the Janzen–Connell hypothesis (Janzen 1970), the optimal defense theory

(Rhoades 1979), the growth–differentiation balance hypothesis (Herms and Mattson 1992), and many others. These values also form the foundation of theories explaining the evolution of plant traits (Coley and Aide 1991) and the formation of biogeographical patterns (Moles et al. 2011). However, the pool of primary data regarding plant losses to insects that has served as the basis for numerous generalizations has never been subjected to rigorous examination to confirm that these data provide unbiased estimates of the amounts of plant biomass consumed by insects. The diversity in methodologies of data collection was recently suggested as one possible reason underlying the lack of consistency among studies exploring the levels of herbivory along environmental gradients (Andrew et al. 2012). In particular, studies may differ in the extent to which various biases have impacted the research process.

For example, the selection of plant species and study sites for measurements of foliar losses to herbivores, as well as of the timing of the measurements, may be affected by the researcher's expectations or preconceptions, which in turn may depend on both the researcher's personal experience and on the hypothesis/theory that the researcher believes to be true. A striking example of this type of effect was revealed in our previous study (Kozlov et al. 2014). For decades, insect herbivory was commonly believed to be highest in tropical regions (Coley and Aide 1991, Coley and Barone 1996, Schowalter and Lowman 1999), and this idea formed the foundation of a number of ecological hypotheses and generalizations (Schemske et al. 2009, Moles et al. 2011). Nevertheless, we found that records made in the tropics had overestimated the community-wide losses of plant foliage to insects due to confirmation bias, i.e., collecting data in a way that confirms the influential theory. In line with this result, many recent studies have found no support for the greatest levels of herbivory in the tropics (Moles et al. 2011, Kozlov et al. 2015a). This example shows the importance of accounting for biases, especially unconscious ones, in the development of scientific knowledge. However, the stages and the specific parts of the research process where biases are most likely to occur are not clear, nor are the consequences of biases for the generalizations based on primary studies. These consequences are especially important due to the increasing use of meta-analysis as a tool for research synthesis in ecological and environmental sciences (Koricheva et al. 2013a, Gurevitch et al. 2018).

The aim of the present study is to understand the extent of the influence of various biases on publications addressing the impacts of different environmental factors that determine spatial patterns in losses of woody plant foliage to insect herbivores, focusing on different stages of scientific research from planning to publication. For this purpose, we conducted a meta-analysis of studies addressing insect herbivory in different environments (habitats, geographical regions, or experimental treatments) and compared the magnitude of the ESs between studies that were likely and unlikely to be influenced by cognitive biases. We asked whether studies based on methods of data collection that are prone to biases tend to overestimate the effects under study. To answer this question, we tested the following specific predictions: (i) studies yield greater ESs when herbivory is measured in a single or a few plant species than in multiple species (i.e., community wide); (ii) publications that formulate the research hypothesis (or clearly state the authors' expectations) report greater ESs than publication that do not formulate the hypothesis and do not state expectations; (iii) the method of selection of plants individuals and individual leaves for the assessment of herbivory affects the magnitude of ESs, with random or blinded selection resulting in the smallest ESs; (iv) awareness of the research hypothesis by the sample collector and measurer of herbivory increases the ESs, whereas blinding of the sample origin during measurements of herbivory reduces of the ESs. We also examined temporal trends in the publication of case studies with potentially biased conclusions, and we asked whether these trends differ between journals with high and low impact factors.

MATERIALS AND METHODS

Database

A huge number of publications is devoted to environmental factors driving spatial patterns of plant losses to insects. We did not perform an exhaustive search of this literature, because we planned to conduct a detailed examination of research methods through communication with the authors of the studies selected for our meta-analysis. Therefore, to make the research feasible, we had to limit the number of studies in our database to slightly over 100. A typical meta-analysis in plant ecology, performed during the past decades, is usually based on several dozens of primary studies (mean 63, median 41; Koricheva and Gurevitch 2014). Thus, the number of studies in our database (MetadataS1, DataS1) is higher than the average number of publications in meta-analyses in the domain of ecological and environmental studies, and we believe that this number is sufficient to uncover multiple sources of variation in spatial patterns of herbivory, including differences in methodology among the primary studies.

As the starting point for this study, we used a database that had been created for the analysis of global patterns in the background losses of woody plant foliage to defoliating insects (selection criteria and the earlier version of the database were published by Kozlov et al. 2015a,b). From this database, which contained 490 publications as of 30 March 2018, we used all publications that explored three major types of environmental factors that potentially affect insect herbivory: habitat disturbance (e.g., fragmentation, pollution, logging, burning, warming, and CO_2 enrichment), geographic variation (e.g., latitude, altitude, and biomes), and the types of natural habitats (e.g., wet vs. dry forest, nutrient-rich vs. nutrient-poor habitats, and more diverse vs. less diverse plant community). The presence of quantitative data on the background insect herbivory in natural ecosystems and the comparison between the two environments were the only criteria for inclusion of the publication in our meta-analysis; therefore, we presume that our choice of studies was not biased.

The 125 selected publications reported mostly the outcomes of observational studies, but eight of them described large-scale field experimental studies that estimated natural herbivory (e.g., enhanced CO_2 level or soil warming).

Collecting information about methods used in primary studies

We inspected the Introduction section of each paper to understand the aim of the study, and we recorded whether explicit research hypotheses or predictions were formulated concerning the differences in herbivory between the compared environments, or whether the authors only searched for differences between the environments. The studies aimed at hypothesis testing were further divided into two groups, based on whether or not the direction of the effect was predicted.

To evaluate the quality of the research methodology and to decide whether the study was prone to biases, we searched each paper for the following information:

1. In how many plots/sites per habitat/treatment (environment hereafter) and on how many plant individuals was the herbivory measured?
2. In how many plant species was the herbivory measured? Was it measured in all plant species in the plot/site (community-wide estimate), or in selected plant species?
3. If herbivory was measured on selected plant species: what was the reason for this selection?
4. For how many years was the herbivory measured?

5. How were plant individuals, branches within a plant (where appropriate), and individual leaves selected for herbivory measurements? We distinguished between the following options: no selection (all available plants or all leaves on the plant were sampled); random selection (some randomization procedure was applied or trees/leaves were haphazardly selected before leaf damage became apparent); systematic selection (e.g., every fourth leaf on the shoot was selected); haphazard selection (selection was made non-blindly and did not involve any specific algorithm aimed at a decrease in subjectivity). Haphazard selection clearly differs from other ways of selection because it can be influenced by researcher's expectations, while random and systematic selection cannot be affected by confirmation bias.
6. Was herbivory measured in situ (without sampling the leaves) or in the laboratory (from collected leaves or their images)?
7. Was herbivory estimated from plant leaves or from their digital images?
8. Was herbivory estimated visually, or with a grid, or by image analysis?
9. Was the person who sampled leaves for measurements of herbivory (or conducted measurements in situ) aware (author) or not aware (technician) of the aim of the study?
10. If herbivory was measured from sampled leaves or their images, were the measurements conducted by a person who was aware or not aware of the aim of the study?
11. If herbivory was measured from sampled leaves or their images, was the information on sample origin available (i.e., did a label include this information?) or not available (i.e., did a label include only the code of the sample?) to the person who conducted the measurements of leaf area lost to insects?

None of 125 publications selected for our meta-analysis, including 10 papers published by ourselves, contained the complete information necessary to answer all these questions and to decide whether the outcome of each individual study may have been affected by different biases. Therefore, we attempted to contact the corresponding authors of all publications. In total, we sent 114 requests (some of which contained questions on more than one paper published by the same author; some requests on the same paper were sent to another co-author if the corresponding author did not respond to our e-mails). The authors of four papers had retired long ago, so we were unable to contact them. Each request contained one to nine questions (median value: five questions). Some authors never responded to our request, so we ultimately obtained complete information on relevant aspects of the research methodology for 103 of the 125 publications. For remaining 22 publications, we used information reported unequivocally in the papers; but when some details of research methodology were not described, we treated this information as missing.

Data extraction and meta-analysis

From each publication, we calculated the Hedges' d measure of the ES; i.e., a normalized difference in herbivory based on means, standard deviations (SD), and sample sizes (n) for the two compared environments. As a rule, we calculated one ES from one paper; but six of 125 papers explored effects of two different factors, so we calculated two ESs from each of these six papers. Therefore, we performed 131 comparisons between the levels of damage of woody plant foliage by insects in different environments (termed 'studies' hereafter).

The majority of the studies selected for our meta-analysis (105 studies) reported the data separately for each of the compared environments. When the primary study used correlation analysis (26 studies), we divided the data collected from the environmental gradients into two groups so that these two groups could be attributed to contrasting environments (e.g., high- and low-altitude sites, high- and low-diversity plots). These two groups of studies yielded effects of similar magnitudes (environmental gradients: d = 0.66; contrasting environments: d = 0.68, respectively; Q_B = 0.04, df = 1, P = 0.85) and were therefore combined in all further analyses.

The studies analyzed by us reported SD as variations between plant species, plant individuals, leaf numbers, study sites, and sampling dates; consequently, sample sizes in these studies reflected different aspects of spatio-temporal variation in insect herbivory. Presentation of data from 55 studies (sometimes complimented by information received from the authors) allowed re-calculation of means, SD, and n so that they reflected variation among plant individuals. For example, when the authors calculated SD as among-year variation and presented the data as means and standard error (SE) or SD reflecting variation among plant individuals for each study year, we calculated the weighted pooled SD value from year-specific SE/SD values using a web-based calculator (https://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/Pooled.htm). For these studies we compared ESs based on both pooled and unpooled SD values and found that the differences in the magnitude of ESs were not significant (pooled: d = 0.69, unpooled: d = 0.64; Q_B = 0.25, df = 1, P = 0.62) as well as for 86 studies based on plant individuals (pooled: d = 0.72, unpooled: d = 0.66, Q_B = 0.18, df = 1, P = 0.68). Across all database, studies based on plant individuals (86 studies) and studies based on study sites (36 studies) yielded ESs of similar magnitudes (Q_B = 0.003, df = 1, P = 0.95). Still, whenever possible, we used sample sizes reflecting the numbers of plant individuals to minimize the variation among studies caused by heterogeneity in SD and n.

In the majority of studies, such as those comparing herbivory in two habitats or correlating herbivory with latitude, classification of the environments as either ‘control’ or ‘treatment’ was not possible. Therefore, in most of our analyses, we considered only the absolute value of ES (i.e., the magnitude of the differences in herbivory). However, when we tested the influence of the unconscious confirmation bias on the outcome of the study, following Holman et al. (2015), we changed the sign of the ES to negative if the reported result opposed the author's predictions on the direction of the effect (14 studies).

We performed our meta-analysis using the random effects categorical models in the MetaWin 2.0 program, assuming that studies differ by sampling error as well as by a random component in the ESs (Rosenberg et al. 2000). The effects were considered statistically significant if the 95% confidence interval of the mean ES (CI_{95}) did not overlap zero. The variation in the ES values within and among the classes of explanatory variables was explored by calculating the heterogeneity indices (Q_T and Q_B respectively) and testing them against χ^2 distribution (Gurevitch and Hedges 2001). Temporal trends in ESs were studied by meta-regression of ES against publication year; temporal changes in sample size were analyzed by calculating the Pearson correlation coefficient between the year of publication and sample size. The proportions of studies that used different methods were compared by frequency analysis (chi-square test; SAS Institute 2009).

RESULTS

Overview of the database

Our meta-analysis is based on 125 papers published from 1977–2018 (MetadataS1, DataS1). We classified the environmental factors explored in the publications included in our meta-analysis into three major groups: habitat type or habitat quality within the same geographical region (e.g., wet vs. dry forests, nutrient-rich vs. nutrient-poor habitats, low diversity vs. high diversity forests, for a total of 56 ESs); habitat disturbances within the same geographical region (e.g., imposed by fragmentation, pollution, logging, burning, warming, or CO_2 enrichment, for a total of 52 ESs); and geographic variation (i.e., differences between habitats located in different latitudes, altitudes, or biomes, for a total of 23 ESs).

Overall variation in herbivory between the environments

Overall differences in the losses of woody plant foliage to insects between the studied environments did not overlap zero ($d = 0.66$, $CI_{95} = 0.56$ to 0.76), indicating that the papers selected for our meta-analysis reported significant spatial variations in herbivory; this variation was highly heterogeneous ($Q_t = 194.1$, df = 130, $P = 0.0002$).

The three major types of environmental factors listed above caused similar effects on herbivory ($d = 0.70$, $d = 0.65$ and $d = 0.67$, respectively; $Q_B = 0.21$, df = 2, $P = 0.90$). Studies conducted in different geographic zones (tropical, subtropical, temperate, or subpolar) yielded similar ESs ($Q_B = 0.90$, df = 3, $P = 0.83$). Observational and experimental studies did not differ in the reported magnitude of the detected effects on herbivory ($d = 0.66$ and $d = 0.66$, respectively; $Q_B = 0.0001$, df = 1, $P = 0.99$). Thus, the discovered heterogeneity is associated with other differences among individual studies, most likely with differences in the methods used by the authors.

Factors affecting the outcome of the study

Numbers of plant species and study years. - Only 17.6% of the studies explored the differences in the community-wide levels of herbivory, as assessed either from abscised leaves or from leaves sampled from all plant species on the plot. More than half of the studies (52.7%) measured herbivory in a single plant species. The most frequent explanation for the selection of one or more particular plant species was that the species was/were common in study area: this reason was mentioned in 94 of 108 (87%) studies based on selected plant species.

Studies in which plant species (either single or several) were selected for measurements of insect herbivory yielded greater ESs than studies where no selection was performed (i.e., where the authors sampled all plant species in the plot and thus obtained community-wide estimates of herbivory) (Fig. 1). The magnitude of the reported differences between the environments was greatly influenced by the number of plant species used to measure insect herbivory: a greater number of species assessed resulted in smaller detected effects (Fig. 1). ESs also tended to be smaller in studies based on two or more years of herbivory measurements than on single-year studies, and primarily in studies of geographic patterns (Fig. 1).

Selection of plants and leaves. - Only about two-thirds (65.2%) of the information on the methods used to select trees, branches, and leaves for measurements of herbivory were reported in the publications used in our meta-analysis. Moreover, clarifications provided by the authors in response to our requests revealed that 26.2% of this information was reported incorrectly. The most common error (44 identified cases) was the use of the term ‘random selection’ in situations when the authors did not apply any randomization procedure, so the selection should therefore have been classified as haphazard.

The review of research methodology associated with selection of plants, branches, and leaves and with measurements of leaf area lost to insects suggested that a substantial fraction of the studies included in our meta-analysis was prone to different kinds of cognitive biases. First, the selection of trees, branches, and/or leaves was haphazard in 53.6% of studies. Second, 82.6% and 81.2% of the leaf sampling and measurement of herbivory, respectively, were generally conducted by the authors, i.e., by the persons who were aware of the hypotheses being tested and/or of the expected research results and were interested in confirming their expectations. Third, only 10.1% of measurements of herbivory were blinded with respect to leaf origin.

The method of selection of plant individuals for measurements of herbivory did not influence the outcome of the study (Fig. 2). By contrast, leaf selection procedure affected the magnitude of the reported differences in herbivory: the greatest differences between the environments were found when leaves were selected haphazardly, and thus their selection may have been unconsciously influenced by the attitudes of the researchers who made this selection. When leaves were not selected (i.e., all leaves on the tree or branch were sampled), or a particular selection protocol (randomization procedure or systematic sampling) was used to avoid the subjectivity in leaf choice, the reported differences were considerably smaller than in the case of haphazard selection (Fig. 2).

Measurement of herbivory. - The differences between the environments in the percentage of leaf area consumed by herbivores did not depend on whether the measurements were conducted in situ or in the laboratory by processing collected leaves or their images ($d = 0.46$ and $d = 0.58$, respectively; $Q_B = 1.06$, $df = 1$, $P = 0.30$). Similarly, visual estimations of leaf damage, the use of grids and image processing yielded similar differences between the environments ($d = 0.47$, $d = 0.55$ and $d = 0.56$, respectively; $Q_B = 0.51$, $df = 2$, $P = 0.\ 77$). However, when sampling of leaves was conducted exclusively by the authors, the reported differences were greater than in the cases, when technicians were involved in sampling (Fig. 2). Moreover, when herbivory was measured in the laboratory, the effect was greater when the measurer was aware of sample origin that when the measurer was blinded to the sample origin (Fig. 2).

Formulation of the predictions. - The majority of publications (75.2%) used in our meta-analysis tested at least one research hypothesis or formulated some prediction(s) about pattern(s) in herbivory; this proportion was greater in journals with an impact factor (IF hereafter) exceeding 1.6 than in journals with lower IF (72.9% and 61.5%, respectively; $\chi^2 = 5.67$, df = 1, P = 0.017). Of all stated predictions, 67% found at least partial support in the studies, and this proportion did not differ between high- and low-impact journals ($\chi^2 = 2.11$, df = 1, P = 0.15). Of 74 studies that predicted the direction of the effect, 14 studies (18.9%) revealed patterns that opposed the research hypotheses or predictions.

The outcome of a study did not depend on prior formulation of the research hypothesis in the paper (formulated: d = 0.48 and not formulated: d = 0.62; $Q_B = 1.13$, df = 1, P = 0.29), nor did it depend on expectations about the direction of the effect under study (direction predicted: d = 0.46, direction not predicted: d = 0.54; $Q_B = 0.24$, df = 1, P = 0.62).

Journal impact factor and publication year. – Across the entire database, the reported differences in herbivory between the environments significantly decreased with the year of publication (meta-regression: Q = 6.09, P = 0.013) and with the increases in a journal's IF (Fig. 3). In particular, 13 of 14 studies that revealed patterns opposing the research hypothesis or predictions were published in journals with IFs exceeding 1.6, and 12 of them were published quite recently (from 2010–2018). The decrease in the magnitude of the ES with time was mostly due to high-impact journals (with IFs exceeding 2.5), while this trend was not significant in other journals (Fig. 3). Sample size did not correlate with the year of publication (r = -0.07, n = 131, P = 0.44).

Methods of leaf selection that are prone to unconscious biases (e.g., haphazard selection) were used twice as frequently in studies published in low-impact journals (IFs below 1.6) than in studies published in journals with higher IF (31.4 and 15.9%, respectively; $\chi^2 = 3.72$, df = 1, P = 0.05). Also the studies published in low-IF journals suffered more frequently from pseudoreplication in terms of the number of study sites or experimental units (33.7% and 61.5%, respectively; $\chi^2 = 8.72$, df = 1, P = 0.003). Across all journals, the percentage of pseudoreplicated studies that used a single site in each habitat decreased from 51.8% in 1977–2008 to 30.1% in 2010–2018 ($\chi^2 = 4.05$, df = 1, P = 0.04). The percentage of these pseudoreplicated studies in high-impact journals dropped during the last years (from 2010–2018) to 19%, compared to 47.5% in other journals ($\chi^2 = 6.48$, df = 1, P = 0.01).

DISCUSSION

Uncovering biases by means of meta-analysis

Psychologists have described a great number of cognitive biases that influence human perception, reasoning, and memory (Forstmeier et al. 2017). Confirmation bias (Rosenthal 1976, Nickerson 1998) is a ubiquitous phenomenon, particularly in science; however, within the biological sciences, confirmation bias has received relatively little attention, and only very recently has the importance of confirmation bias been recognized for ecological and environmental research (Kozlov et al. 2014, Kozlov and Zvereva 2015, Holman et al. 2015, Kardish et al. 2015). Avoidance of confirmation bias requires that experimenters or observers blind themselves by concealing any information about research hypotheses or the treatment conditions of a specific sample (Noseworthy et al. 1994, van Wilgenburg and Elgar 2013, Kardish et al. 2015). Therefore, confirmation bias is usually demonstrated by comparing blind and non-blind studies (Kozlov et al. 2014, Tuyttens et al. 2014, Kozlov and Zvereva 2015), and meta-analyses that compare these two kinds of studies both across and within disciplines have shown a tendency toward higher ESs in non-blind than in blind studies (van Wilgenburg and Elgar 2013, Holman et al. 2015, Saltaji et al. 2018). Our meta-analysis of spatial patterns in herbivory not only confirmed this tendency, but it also uncovered other sources of cognitive biases in biological research (Table 1). In general, the use of research methods that are prone to different biases results in overestimation of the effects under study and, consequently, leads to overestimation of the average ESs in subsequent quantitative research syntheses (Table 1).

Biases associated with selection of plant species and duration of the study

The meta-analyses conducted in the field of ecology clearly demonstrated a disproportional amount of information on certain ecosystems, high-rank taxa, and individual species (Gurevich and Hedges 1999). This 'biased sampling of the natural world' (Jennions et al. 2013) may lead to wrong conclusions of meta-analyses due to over-representation of some study systems. In some research areas, the studies may be concentrated on just one or a few 'model' species, and this may lead to completely wrong fundamental conclusions. One striking example concerns studies of gene expression that have drawn conclusions about principal differences between plants and animals just because the model plant, *Arabidopsis thaliana*, appeared exceptional among plants in the studied character (Lloyd and Davies 2013). In the area of the spatial patterns in herbivory almost no studies (and none of the studies used in our meta-analysis) were performed on gymnosperms, despite the ecological and economic importance of gymnosperms in many regions. This deficiency arose presumably due to practical difficulties associated with measurements of herbivory in needle-bearing trees. Therefore, the detected patterns of geographic variation in herbivory (e.g., by Kozlov et al. 2015a, b) may only be valid for woody angiosperms and not for all woody plants.

One frequent example of non-random data collection is selection of species that occur at high densities (Jennions et al. 2013). Indeed, in studies included in our meta-analysis, most plants were selected for measurements of herbivory based on their high abundance, as declared in the publications. However, our database shows no systematic bias toward a certain species because of the great variety of habitats and geographic zones, each of which differ in the predominating plant species in the community. Conversely, a possibility remains that foliar losses to insects systematically differ in terms of responses to environmental factors when comparing abundant versus rare plant species. For example, more abundant plants may suffer higher levels of herbivory due to the higher probability that these species will be found by herbivores, as predicted by the resource availability hypothesis (Endara and Coley 2011).

In an earlier study (Kozlov et al. 2014), we found that the reported foliar losses to herbivory in Brazil were two-fold higher in studies involving 1–3 species of woody plants than in studies involving more than 10 species. In the current study, we also demonstrated that differences in herbivory between the environments decreased with the number of plant species used for herbivory measurements, and were lowest when the study involved more than 10 plant species. Among these studies, 80% measured herbivory from all plant species growing in a study site. The community-wide estimates of herbivory (or their approximations based on measurements of herbivory from multiple plant species) clearly provide a more adequate reflection of the differences between the environments when compared to results from studies based on a few selected plant species. Consequently, the outcomes of multi-species studies are most important for understanding the impacts of environmental factors on herbivory, for uncovering global patterns in herbivory, and for making predictions on the roles of herbivores in ecosystem functioning. Therefore, we conclude that studies conducted on a single or few plant species not only provide overestimated levels of herbivory (Kozlov et al. 2014), but they also generally overestimate the magnitude of spatial variation in plant losses to insects.

Plant species differ considerably in the levels of herbivory they experience (Lowman 1995, Brenes-Arguedas et al. 2008, Ruiz-Guerra et al. 2010). The differences in species-specific values of herbivory between few-species and multi-species studies can be explained by unconscious avoidance by the researchers of plant species that showed no or little foliar damage when using non-blind methods of species selection (Kozlov et al. 2014). Only rarely (one study in our database) did the authors acknowledge that a study species was selected because it showed greater damage from insects when compared to other plant species (Peter et al. 2015).

The decrease in the magnitude of the responses to environmental factors with an increase in the number of plant species demonstrated in our meta-analysis is more difficult to explain. Nevertheless, we suggest that this pattern may arise for two reasons. First, plant species display diverse responses to environmental factors. When a given species grows in two different environments (e.g., in dry and wet forests or in disturbed and undisturbed sites), one of the habitats under study will always be less favorable for that species than another habitat. This difference may affect plant losses to insects because some insect species inflict higher damage on more stressed plants (White 1974), whereas other insect species prefer more vigorous plants (Price 1991). In community-wide studies, the responses of different plant species to the same environmental factors are likely to differ not only in their magnitude, but also in their direction, resulting in variable consequences for plant-feeding insects. When combined, these responses are likely to yield low or even zero overall effects on plant losses to insects. This pattern was found in earlier meta-analysis of the effects of herbivory on the leaf life span of woody plants, where stronger effects were observed in studies that included a single herbivore species than in studies of multi-species insect assemblages (Zvereva and Kozlov 2014).

However, if the selection of species in single-species studies was random, the effect size in the meta-analysis would not differ between single- and multi-species studies. Therefore, we suggest that greater differences in herbivory between environments will be observed in studies in which species are selected than in multi-species studies due to researcher bias. This bias is defined by Gurevitch and

Hedges (1999) as a tendency to collect data on organisms or under conditions in which one has reasonable expectations of detecting statistically significant effects. This kind of bias is not necessarily unconscious, because this selection could be made deliberately to increase probability of obtaining statistically significant results. In this case, the authors might select the most responsive plant species, based on their previous experience or on earlier studies of other researchers. The high occurrence of this bias is indirectly confirmed by the fact that only one paper from our database (Nuckols and Connor 1995) mentioned that the plant species were selected without knowledge of herbivore attacks on them in either of the compared habitats.

The differences observed between single-species and multi-species studies may also be due to selective reporting. We have noticed that when herbivory on various studied plant species differs in response to environmental factors, the authors present quantitative data (which can be used in meta-analysis) only for the species that fit the research hypothesis, while ignoring or reporting incompletely data for species that do not fit the research hypothesis (e.g., Knepp et al. 2005, Sobek et al. 2009). This selective reporting is common in many scientific disciplines, including ecology and evolution (reviewed by Parker et al. 2016). Moreover, the heterogeneity of the results is often not mentioned in either the discussion or the abstract; for example, Knepp et al. (2005) found support for their hypothesis only in one of three study years and for only a few of the 12 plant species selected; however, they concluded that their study, in general, supported the hypothesis.

Similarly to the effects of the study species, the detected environmental effects tended to decrease when observations or experiments extended beyond a single year. This difference may be explained by among-year variation in the responses of herbivory to the environmental factors under study, which was demonstrated by several studies from our database (Bogacheva 1986, Knepp et al. 2005, Kozlov 2015). We found especially large differences between single-year and multi-year studies addressing geographic variation in herbivory. Changes in herbivory with latitude or altitude may differ among years not only in magnitude, but even in their direction (Kozlov et al. 2013), indicating considerable interaction between geographic pattern and current-year weather conditions. A striking example is provided by the study of sap-feeding insects: their load on plants decreased with latitude in typical summers but increased in an exceptionally hot summer and was independent of latitude during a warm summer (Kozlov et al. 2015c).

Higher magnitudes of the effects detected for single-year studies compared to multi-year studies may also be explained by selective reporting. Quite possibly, the results from the year(s) that did not fit the research hypothesis were simply not included in some papers. However, this kind of selective reporting, when some results go completely unreported, is hard to uncover (Parker et al. 2016).

In some situations, the selection of the study year may also be biased. For example, the meta-analysis of the impact of point polluters on terrestrial insects showed that insect herbivore density was preferentially measured in the years when plant damage in polluted sites was apparent (Zvereva and Kozlov 2010), i.e. the pattern fitted the influential hypothesis (Führer 1985, Baltensweiler 1985) on the increase of herbivory under pollution impact. This bias resulted in overestimation of the pollution effects on insect herbivory. However, in the majority of the studies, the selection of study year was determined by factors other than levels of herbivory and thus can be considered as random. Nevertheless, the spatial pattern observed in a single-year study may be transient or accidental and may not agree with the general trend. This calls for conducting field studies for at least two years before a justified conclusion about a spatial pattern in herbivory can be made.

In summary, single-species studies conducted over a single year tend to overestimate the magnitudes of the differences in plant losses to insects between the environments, and these results are less likely to reflect the general patterns in herbivory than are multi-species and multi-year studies.

Biases associated with leaf sampling and with measurements of herbivory

Sampling procedure can considerably influence the conclusions of ecological studies (Albert et al. 2010, Mörsdorf et al. 2015). In our meta-analysis, the impact of a researcher's expectations on the outcomes of the study is clearly seen from the greatest reported effects for leaves that were sampled by authors only, while an involvement of technicians (who are usually not aware of the hypothesis or have less interest in confirming it) decreased the magnitude of the differences in herbivory between the environments. This bias was further detected in the next stage of sample processing in the laboratory: the effect was higher when the measurer was aware of the sample origin. Thus, when the person who selects leaves and measures herbivory is aware of the study predictions and is interested in confirming these predictions (i.e., the research hypothesis is not blinded), and when that person is aware of the sample origin (i.e., the treatment conditions are not blinded), the study is likely to overestimate the impacts of environmental factors on levels of herbivory due to confirmation bias. This result is in line other meta-analyses within medical (Saltaji et al. 2018) and biological (van Wilgenburg and Elgar 2013, Holman et al. 2015) sciences, in which non-blind studies generally yield higher effect sizes when compared with studies using blinding protocols.

By contrast, we found similar ESs for studies where plant losses to insects were estimated visually or by using relatively sophisticated (e.g., digital images) or less high-tech (e.g., grids) methods and devices. We could expect that, in a meta-analysis, the visual estimates of leaf damage, which are presumably less precise and accurate than other methods, would result in smaller ESs due to larger SDs; but this expectation was not met. This result is in line with studies showing that visual estimates of herbivory are as accurate as image processing (Kozlov and Zvereva 2018 and references therein), and it indicates that when measurements are not blinded, the outcomes of these measurements are much more influenced by 'wishful thinking' of the measurer than by the particular method used.

However, we discovered that factors other than lack of blinding may also lead to overestimation of the effects. We found, on average, 72% greater effects for studies measuring herbivory from leaves haphazardly selected by the researcher when compared to studies that measured herbivory from all leaves of the plant or its branch (i.e., performing no selection of individual leaves), from leaves selected before herbivory was apparent (blind selection), from leaves sampled systematically (e.g., every fourth leaf on branch, starting from its tip, was sampled), or from leaves sampled randomly (in the strict sense of this word, where some randomization procedure was applied) . This indicates that when a researcher has the possibility of selecting leaves for estimations of herbivory, this selection introduces considerable bias and leads to overestimation of the differences in herbivory between the environments. This unconscious bias is likely caused by certain expectations or motivations to confirm the research hypothesis, and can therefore be classified as confirmation bias.

Interestingly, we did not find any bias in the selection of plant individuals for sampling. In contrast to leaf sampling, studies that sampled plant individuals haphazardly (i.e. were prone to confirmation bias) yielded similar effect size to those obtained with studies that selected plants before herbivory was apparent or that used some randomization procedure or that sampled all plants in a plot. This difference hints that it is generally difficult to unconsciously estimate plant-wide level of herbivory when selecting a tree for sampling, especially a large one (and large trees were used in the majority of the studies included in our database), whereas during the leaf sampling, the collector can unconsciously assess the level of damage of each individual leaf

Contrary to our expectations, the formulation of a research hypothesis appeared to be not necessary for the emergence of an unconscious confirmation bias: our meta-analysis did not find any differences in ESs between groups of studies that strictly stated or did not state any hypothesis and the aim of the study was just to compare herbivory in two types of habitats. We expected to detect this difference not only because formulation of the hypothesis creates a strong motivation to confirm it, but also because, in many studies, the hypothesis is formulated after getting the result ('HARKing' – Hypothesizing After Results are Known) (Kerr 1998). The latter should lead to higher ESs in studies that formulated a hypothesis because if the study finds strong effect, the authors would tend to HARK more frequently than if the study does not find an effect or finds a weak one. However, this did not appear to be the case. Our result is in line with the opinion of Forstmeier et al. (2017) that a priori hypothesis testing or HARKing does not substantially differ in terms of the likelihood of a positive finding. We suggest that when a study compares habitats or treatments, researchers always search for differences, both consciously and unconsciously, and this is sufficient to trigger confirmation bias. The formulation of a hypothesis in the publication is more a matter of journal requirements (see below).

Our previous study (Kozlov et al. 2014), which compared blind and non-blind sampling, showed that estimates of plant damage by insect herbivores in the tropics were considerably influenced by unconscious selection of plants with higher-than-average levels of herbivory, and that this bias could lead to overestimation of community-wide losses of plant foliage to insects. Our current meta-analysis demonstrated that differences in herbivory between the environments may also be overestimated due to unconscious biases. Ecological studies involving the assessment of herbivory may appear especially prone to these biases because herbivore damage may be perceived visually during sampling, and these biases may affect several stages of the study: selection of species, plant individuals, branches of the plant (on woody plants), leaves for measurements, and the measurement process itself. Based on the analysis of published papers and the responses of authors to our questions, we concluded that all 125 publications included in our database have potentially been affected by unconscious confirmation bias at one or several stages of the research. This is much more than the 50.4% reported in other domains within ecological science, as found by Kardish et al. (2015) who analyzed 248 publications in 11 high-impact journals. We showed that these multiple possibilities of the occurrence of unconscious bias during assessments of herbivory led not only to overestimation of herbivory levels, but also to overestimation of the magnitude of the effects of different environmental factors on spatial patterns in herbivory.

Unconscious biases in studies of insect herbivory may be avoided by different methods. First, selection procedures should be avoided (e.g., by sampling all plants on the plot and all leaves on the plant, or by selecting of plants and/or leaves before herbivore damage becomes apparent). Another way is to apply a true randomization procedure or to sample plants/leaves systematically. However, even when the sampling of plants and leaves is unbiased, confirmation bias may still affect the outcomes of the study during the measurements of leaf losses to insects. This bias is unavoidable when herbivory is estimated in situ, but when leaves (or their images) are collected and analyzed in the laboratory, confirmation bias may be avoided by blinding the samples (i.e., labeling them in a way that does not provide a measurer with any information about the sample origin). However, only three papers (2.4%) in our database mentioned blinding of samples, while in the database of Kardish et al. (2015), this percentage was 13.3%. However, the authors of 11 additional studies reported blinding of sample origin when responding to our direct question. This result indicates that the authors clearly underestimate the importance of blinding and it justifies the necessity of requesting authors for details of their methods.

Furthermore, it suggests that the values provided by Kardish et al. (2015), which were based on the published information only, underestimate the real frequency of the use of blind methods. In any case, studies including an assessment of herbivory seem to use blind methods more rarely when compared to studies from other research areas within the domain of ecological, evolutionary, and behavioral research.

Publication bias and temporal trends

The number of hypotheses being tested per paper is increasing in ecology (Low-Décarie et al. 2014). In our database, 76.3% of studies erected one or more research hypotheses or formulated one or more predictions, and this proportion was greater in journals with higher IFs. The percentage of published studies that support the hypothesis in our database (67%) was quite close to the percentage of studies detected by Finelli (2010) for the research domain ‘Environment and Ecology’ (74%). In an earlier paper (Csada et al. 1996), only 8.6% of studies were found to present non-significant tests of their main hypothesis in a cross-section of biological journals. This difference may either reflect temporal trends (see below), or it may indicate that our study domain (Environment and Ecology) is suffering less from publication bias than are biological sciences in general, which is in line with the comparative data presented by Fanelli (2010). Moreover, although the proportion of studies in our database that tested research hypotheses is significantly greater in journals with high and medium impact factor than in low-impact journals, the frequency of publications supporting the research hypothesis does not depend on a journal's IF. Contrary to expectations on preferential publication of strong significant effects in high-ranked journals (Leimu and Koricheva 2004), the mean magnitude of the effects detected by our meta-analysis was smaller for papers published in these journals than in journals with lower IFs. This result confirms the negligible impact of publication bias on the studies included in our database. Significant changes in the magnitude of research findings over time have been reported in many meta-analyses from different areas of ecology, and the majority of studies discovered a decrease in the ESs with publication year (summarized by Koricheva et al. 2013b). This trend may have several explanations. First, it can be due to the smaller sample sizes used in older studies (Ioannidis 2008). However, we have not found any temporal trends in sample size in our study, in line with the conclusions by Koricheva et al. (2013b) based on 54 meta-analytical studies in ecology and evolution. Jennions and Møller (2002) suggested that the decrease in ESs with time results from delayed publication of studies reporting small or non-significant effects; however, evidence for time-lag bias was not found in ecological studies (Jennions et al. 2013). Another explanation lies in the development of the evidence for any ecological hypothesis. This process starts from strong supportive evidence of a newly formulated hypothesis, but later on accumulation of disconfirming evidence leads to a decrease in the ESs with time (Leimu and Koricheva 2004). However, this explanation is not applicable to our database, because different papers test different hypotheses that have different ‘dates of birth’ and different histories of their development. Temporal changes may also be caused by changed in the preference toward particular study organisms and systems (e.g. Nykänen and Koricheva 2004, Saikkonen et al. 2006). In our case, this trend could be due to temporal shifts in the proportion of multi-species studies, which yield smaller effect sizes than do single species studies; however, in our study, the proportion of multi-species studies did not change with time.

Lastly, temporal trends in ESs may result from changes in research methodology. A number of studies demonstrated that improvement in methods, including statistical analysis, resulted in a decrease in the magnitude of the effect (reviewed by Koricheva et al. 2013b). In our opinion, this is most plausible explanation for the temporal trends observed in our study, because a significant trend was detected only in high-impact journals, while the decrease was weak and non-significant in other journals. Stronger demands with respect to research methodology in high-impact journals are confirmed in our meta-analysis by a two-fold less frequent use of methods that are prone to unconscious confirmation bias (haphazard selection of plants or leaves in the first line) when compared to other methods that are not prone to this bias. High-impact journals also publish more multi-species studies, which reflect habitat-specific levels of herbivory more adequately when compared with studies exploring herbivory on a single or few plant species. Moreover, the frequency of pseudoreplication (expressed primarily in the use of a single site in each of the compared environments) is significantly lower in high-impact journals than in other journals, and this difference has especially increased during last decade. In high-impact journals, the ES for environmental effects reported in the publications from 2010-2018 were as low as 0.3, which indicates a high proportion of papers that either did not detect significant effects or that detected effects contradicting the research hypothesis.

The detected temporal trends in the improvement of methodology may at least partly result from Transparency and Openness Promotion (TOP)—a framework that emerged in 2015 and is currently supported by many high-impact journals, with many TOP guidelines that request or require more thorough reporting of methods and results (Parker et al. 2016). Our meta-analysis demonstrated that progress in the quality, presentation, and transparency of methodology has already led to a decrease in the frequencies of methodological flaws and, as a result, to a decrease in the magnitude of environmental effects on herbivory measured during last decade. This decrease in the magnitude of the effect with time, and even an increase in the number of publications that oppose the existing hypotheses (12 of 14 these papers were published after 2010), indicates not only an improvement of methodology, but also a considerable decrease in publication bias against non-confirming results in high-impact journals. A similar shift in publication bias has been demonstrated by Low-Décarie et al. (2014) for three ecological journals. From these patterns, we conclude that studies published in the 1960s−1990s considerably overestimated the magnitude of environmental effects on insect herbivory.

Our results also oppose the opinion that publication bias is especially strong in high-ranked journals, which prefer to publish studies reporting large, highly significant effects (Murtaugh 2002) and/or support a prevailing hypothesis over those that reject it (Leimu and Koricheva 2004). We found that this kind of publication bias is weaker in high-impact journals than in journals with lower ranking, especially in recent years. A similar trend has been demonstrated in the meta-analyses of the impact of pollution on arthropods, plant growth, and plant diversity, where ESs were lower in high-impact publications (Zvereva et al. 2008, 2010, Zvereva and Kozlov 2010). This means that, in our research domain (spatial patterns in herbivory), the pressure on researchers to publish only strong and supportive results is relatively low and steadily decreasing. We conclude that the current trends in herbivory studies overcome the overestimation of the effects due to different biases and thereby improve our knowledge, mostly due to the changing editorial policy of high-impact journals.

CONCLUSIONS

Using the studies of spatial patterns in herbivory as an example, we showed that the properties of human mind may lead to overestimation of the magnitudes of the effects recorded in the course of ecological and environmental studies (Table 1). The magnitude of this overestimation is of high ecological significance: methods that are not prone to biases and thus reflect the real situation more adequately yield small to medium ESs (0.2 and 0.5 according to Cohen 1988), while methods that are prone to biases generally yield large ESs (0.8). Importantly, factors related to methodology appeared the only significant sources of variation in magnitudes of spatial patterns in insect herbivory, and most of this variation likely emerged due to various biases. Whether a hypothesis is formulated or not, researchers always have some expectations concerning the studied patterns and processes, and these expectations trigger unconscious psychological processes, unavoidably biasing the outcome of the study. The biases are especially frequent in studies based on sampling of data from natural ecosystems, and a number of measures should be taken to avoid these biases (Table 1). Selection of study species may be prone to bias and should therefore be avoided when comparing different environments—for example, by sampling all species on the plot. At lower hierarchical levels, when the selection of study units (such as individual leaves within a large tree) is unavoidable, sampling should be randomized in the strict sense (i.e., by applying randomization procedures). Finally, the studies should be blinded with respect to the research hypothesis and sample origin. The information about all these measures taken to avoid biases should be explicitly provided in the publications. Researchers must remain aware of the existence of unconscious biases that occur at different stages of the research and that can considerably influence the magnitudes of the effects under study. Accumulation of biased results in primary studies may ultimately lead to false general conclusions when the outcomes of these studies are summarized in research syntheses.

ACKNOWLEDGEMENTS

We are grateful to the authors of the publications included in our analysis for answering our questions on research methods and for providing additional information, including primary data, which was necessary for calculation of effect sizes. We thank N. G. Yoccoz and anonymous reviewer for their helpful comments to the earlier draft of the manuscript. The study was supported by the Kone Foundation and by the Academy of Finland (projects 276671, 288213 and 316182).

LITERATURE CITED

Albert, C. H., N. G. Yoccoz, T. C. Edwards Jr, C. H. Graham, N. E. Zimmermann, and W. Thuiller. 2010. Sampling in ecology and evolution − bridging the gap between theory and practice. Ecography 33:1028-1037.

Andrew, N. R., I. R. Roberts, and S. J. Hill. 2012. Insect herbivory along environmental gradients. Open Journal of Ecology 2:202-213.

Baltensweiler, W. 1985. Waldsterben − forest pests and air pollution. Journal of Applied Entomology 99:77-85.

Bogacheva, I. A. 1986. Dependence of insect herbivore abundance on the level of fluorine contamination in forest communities. Pages 43-48 *in* B. P. Popov, and O. F. Sadykov, editors. Human-created factors and animal organism: field observations and experiments. Ural Scientific Center, Academy of Sciences of the USSR, Sverdlovsk, Russia (in Russian).

Brenes-Arguedas, T., P. D. Coley, and T. A. Kursar. 2008. Divergence and diversity in the defensive ecology of *Inga* at two Neotropical sites. Journal of Ecology 96:127-135.

Cohen, J. 1988. Statistical power analysis for the behavioral sciences. Academic Press, New York, USA.

Coley, P. D., and T. M. Aide. 1991. Comparison of herbivory and plant defenses in temperate and tropical broad-leaved forests. Pages 25-49 *in* P. W. Price, T. M. Lewinsohn, G. W. Fernandes, and

W. W. Benson, editors. Plant-animal interactions: evolutionary ecology in tropical and temperate regions. Wiley, New York, USA.

Coley, P. D., and J. A. Barone. 1996. Herbivory and plant defenses in tropical forests. Annual Review of Ecology and Systematics 27:305-335.

Csada, R. D., P. C. James, and R. H. M. Espie. 1996. The "file drawer problem" of non-significant results: does it apply to biological research? Oikos 76:591-593.

Endara, M. J., and P. D. Coley. 2011. The resource availability hypothesis revisited: a meta-analysis. Functional Ecology 25:389-398.

Fanelli, D. 2010. "Positive" results increase down the hierarchy of the sciences. PLoS One, 5.

Fanelli, D. 2012. Negative results are disappearing from most disciplines and countries. Scientometrics 90:891-904.

Fanelli, D., R. Costas, and J. P. A. Ioannidis. 2017. Meta-assessment of bias in science. Proceedings of the National Academy of Sciences of the United States of America 114:3714-3719.

Forstmeier, W., E. J. Wagenmakers, and T. H. Parker. 2017. Detecting and avoiding likely false-positive findings − a practical guide. Biological Reviews 92:1941-1968.

Führer, E. 1985. Air pollution and the incidence of forest insect problems. Journal of Applied Entomology 99:371-377.

Gluud, L. L. 2006. Bias in clinical intervention research. American Journal of Epidemiology 163:493-501.

Gurevitch, J., and L. V. Hedges. 1999. Statistical issues in ecological meta-analyses. Ecology 80:1142-1149.

Gurevitch, J., and L. V. Hedges. 2001. Metaanalysis. Combining the results of independent experiments. Pages 347-369 *in* S. M. Schneider, and G. Gurevitch, editors. Design and analysis of ecological experiments. Oxford University Press. Oxford, USA.

Gurevitch J., J. Koricheva, S. Nakagawa, and G. Stewart. 2018. Meta-analysis and the science of research synthesis. Nature 555:175-182.

Hairston, N. G., F. E. Smith, and L. B. Slobodkin. 1960. Community structure, population control, and competition. American Naturalist 94:421-425.

Hergovich, A., R. Schott, and C. Burger. 2010. Biased evaluation of abstracts depending on topic and conclusion: further evidence of a confirmation bias within scientific psychology. Current Psychology 29:188-209.

Herms, D. A., and W. J. Mattson. 1992. The dilemma of plants: to grow or defend. Quarterly Review of Biology 67:283-335.

Holman, L., M. L. Head, R. Lanfear, and M. D. Jennions. 2015. Evidence of experimental bias in the life sciences: why we need blind data recording. PLoS Biology 13(7):e1002190.

Ioannidis, J. P. A., N. Patsopoulos, and H. Rothstein. 2008. Reasons or excuses for avoiding meta-analysis in forest plots. British Medical Journal 336:1413-1415.

Jamieson, M. A., A. M. Trowbridge, K. F. Raffa, and R. L. Lindroth. 2012. Consequences of climate warming and altered precipitation patterns for plant-insect and multitrophic interactions. Plant Physiology 160:1719-1727.

Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. American Naturalist 104:501-528.

Jennions, M. D., and A. P. Møller. 2002. Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. Proceedings of the Royal Society B-Biological Sciences 269:43-48.

Jennions, M. D., C. J. Lortie, M. S. Rosenberg, and H. R. Rothstein. 2013. Publication and related biases. Pages 207-236 *in* J. Koricheva, J. Gurevitch, and K. Mengersen, editors. Handbook of meta-analysis in ecology and evolution. Prinston University Press, Prinston & Oxford, USA.

Kardish, M. R., U. G. Mueller, S. Amador-Vargas, E. I. Dietrich, R. Ma, B. Barrett, and C.-C. Fang. 2015. Blind trust in unblinded observation in Ecology, Evolution, and Behavior. Frontiers in Ecology and Evolution 3:51.

Kerr, N. L. 1998. HARKing: hypothesizing after the results are known. Personality and Social Psychology Review 2:196-217.

Knepp, R. G., J. G. Hamilton, J. E. Mohan, A. R. Zangerl, M. R. Berenbaum, and E. H. DeLucia. 2005. Elevated CO_2 reduces leaf damage by insect herbivores in a forest community. New Phytologist 167:207-218.

Koricheva, J., J. Gurevitch, and K. Mengersen. 2013a. Handbook of meta-analysis in ecology and evolution. Princeton University Press, Princeton & Oxford, USA.

Koricheva, J., M. D. Jennions, and J. Lau. 2013b. Temporal trends in effect sizes: causes, detection, and implications. Pages 237-254 *in* J. Koricheva, J. Gurevitch, and K. Mengersen, editors. Handbook of meta-analysis in ecology and evolution. Princeton University Press, Princeton & Oxford, USA.

Koricheva, J., and J. Gurevitch. 2014. Uses and misuses of meta-analysis in plant ecology. Journal of Ecology 102:828-844.

Kozlov, M. V. 2015. Insect herbivory on two willow species in northern Europe is independent of pollution load. Boreal Environment Research 20:423-430.

Kozlov, M. V., and E. L. Zvereva. 2009. Research, publication, and dissemination biases in observational studies of pollution impact on terrestrial biota. Pages 23-26 *in* K. Vakkilainen & V. Pukkila, editors. Proceedings of FCES-09 Finnish Conference of Environmental Science. 14-15 May 2009, Lahti, Finland. University of Helsinki, Helsinki, Finland.

Kozlov, M. V., and E. L. Zvereva. 2015. Confirmation bias in studies of fluctuating asymmetry. Ecological Indicators 57:293-297.

Kozlov, M. V., and E. L. Zvereva. 2018. Background insect herbivory: impacts, patterns and methodology. Progress in Botany 79:313-356.

Kozlov, M. V., E. J. van Nieukerken, V. Zverev, and E. L. Zvereva. 2013. Abundance and diversity of birch-feeding leafminers along latitudinal gradients in northern Europe. Ecography 36:1138-1149.

Kozlov, M. V., V. Zverev, and E. L. Zvereva. 2014. Confirmation bias leads to overestimation of losses of woody plant foliage to insect herbivores in tropical regions. PeerJ 2:e709.

Kozlov, M. V., V. Lanta, V. Zverev, and E. L. Zvereva. 2015a. Global patterns in background losses of woody plant foliage to insects. Global Ecology and Biogeography 24:1126-1135.

Kozlov, M. V., V. Lanta, V. Zverev, and E. L. Zvereva. 2015b. Background losses of woody plant foliage to insects show variable relationships with plant functional traits across the globe. Journal of Ecology 103:1519-1528.

Kozlov, M. V., A. V. Stekolshchikov, G. Soderman, E. S. Labina, V. Zverev, and E. L. Zvereva. 2015c. Sap-feeding insects on forest trees along latitudinal gradients in northern Europe: a climate-driven patterns. Global Change Biology 21:106-116.

Leedy, P. D., and J. E. Ormrod. 2018. Practical research: planning and design, 12 edition, Pearson, Boston, USA.

Leimu, R., and J. Koricheva. 2004. Cumulative meta-analysis: a new tool for detection of temporal trends and publication bias in ecology. Proceedings of the Royal Society B-Biological Sciences 271:1961-1966.

Lloyd, J. P. B., and B. Davies. 2013. SMG1 is an ancient nonsense-mediated mRNA decay effector. Plant Journal 76:800-810.

Low-Décarie, E., C. Chivers, and M. Granados. 2014. Rising complexity and falling explanatory power in ecology. Frontiers in Ecology and the Environment 12:412-418.

Lowman, M. D. 1995. Herbivory in Australian forests – a comparison of dry sclerophyll and rain forest canopies. Proceedings of the Linnean Society of New South Wales 115:77-88.

Marsh, D. M., and T. J. Hanlon. 2007. Seeing what we want to see: confirmation bias in animal behavior research. Ethology 113:1089-1098.

Moles, A. T., S. P. Bonser, A. G. B. Poore, I. R. Wallis, and W. J. Foley. 2011. Assessing the evidence for latitudinal gradients in plant defence and herbivory. Functional Ecology 25:380-388.

Murtaugh, P. A. 2002. Journal quality, effect size, and publication bias in meta-analysis. Ecology 83:1162-1166.

Mörsdorf, M. A., V. T. Ravolainen, L. E. Støvern, N. G. Yoccoz, I. S. Jónsdóttir, and K. A. Bråthen. 2015. Definition of sampling units begets conclusions in ecology: the case of habitats for plant communities. PeerJ 3:e815.

Nickerson, R. S. 1998. Confirmation bias: a ubiquitous phenomenon in many guises. Review of General Psychology 2:175-220.

Noseworthy, J. H., G. C. Ebers, M. K. Vandervoort, R. E. Farquhar, E. Yetisir, and R. Roberts. 1994. The impact of blinding on the results of a randomized, placebo-controlled multiple-sclerosis clinical trial. Neurology 44:16-20.

Nuckols, M. S., and E. F. Connor. 1995. Do trees in urban or ornamental plantings receive more damage by insects than trees in natural forests? Ecological Entomology 20:253-260.

Nykänen, H., and J. Koricheva. 2004. Damage-induced changes in woody plants and their effects on insect herbivore performance: a meta-analysis. Oikos 104:247-268.

Oksanen, L., S. D. Fretwell, J. Arruda, and P. Niemelä. 1981. Exploitation ecosystems in gradients of primary productivity. American Naturalist 118:240-261.

Pannucci, C. J., and E. G. Wilkins. 2010. Identifying and avoiding bias in research. Plastic and Reconstructive Surgery 126:619-625.

Parker, T. H., W. Forstmeier, J. Koricheva, F. Fidler, J. D. Hadfield, Y. E. Chee, C. D. Kelly, J. Gurevitch, and S. Nakagawa. 2016. Transparency in ecology and evolution: real problems, real solutions. Trends in Ecology & Evolution 31:711-719.

Peter, F., D. G. Berens, G. R. Grieve, and N. Farwig. 2015. Forest fragmentation drives the loss of insectivorous birds and an associated increase in herbivory. Biotropica 47:626-635.

Polis, G. A. 1999. Why are parts of the world green? Multiple factors control productivity and the distribution of biomass. Oikos 86:3-15.

Price, P. W. 1991. The plant vigor hypothesis and herbivore attack. Oikos 62:244-251.

Rhoades, D. F. 1979. Evolution of plant chemical defence against herbivores. Pages 3-54 *in* G. A. Rozenthal, and D. H. Janzen, editors. Herbivores: their interactions with secondary plant metabolites. Academic Press, New York, USA.

Rosenberg, M. S., D. C. Adams, and J. Gurevitch 2000. MetaWin: statistical software for meta-analysis. Version 2.0., Sinauer, Sunderland, Massachusetts, USA.

Rosenthal, R. 1976. Experimenter effects in behavioural research. John Wiley, New York, USA.

Ruiz-Guerra, B., R. Guevara, N. A. Mariano, and R. Dirzo. 2010. Insect herbivory declines with forest fragmentation and covaries with plant regeneration mode: evidence from a Mexican tropical rain forest. Oikos 119:317-325.

Saikkonen, K., P. Lehtonen, M. Helander, J. Koricheva, and S. H. Faeth. 2006. Model systems in ecology: dissecting the endophyte-grass literature. Trends in Plant Science 11:428-433.

Saltaji, H., S. Armijo-Olivo, G. G. Cummings, M. Amin, B. R. da Costa, and C. Flores-Mir. 2018. Influence of blinding on treatment effect size estimate in randomized controlled trials of oral health interventions. BMC Medical Research Methodology 18:42.

SAS Institute 2009. SAS/Stat. User's guide, version 9.2. SAS Institute, Cary, North Carolina, USA.

Schemske, D. W., G. G. Mittelbach, H. V. Cornell, J. M. Sobel, and K. Roy. 2009. Is there a latitudinal gradient in the importance of biotic interactions? Annual Review of Ecology, Evolution and Systematics 40:245-269.

Schowalter, T. D., and M. D. Lowman. 1999. Forest herbivory: insects. Pages 253-269 *in* L. R. Walker, editor. Ecosystems of the world 16: ecosystems of disturbed ground. Elsevier, Amsterdam, Netherlands.

Sobek, S., Ch. Scherber, I. Steffan-Dewenter, and T. Tscharntke. 2009. Sapling herbivory, invertebrate herbivores and predators across a natural tree diversity gradient in Germany's largest connected deciduous forest. Oecologia 160:279-288.

Traniello, J. F. A., and T. C. M. Bakker. 2015. Minimizing observer bias in behavioral research: blinded methods reporting requirements for behavioral ecology and sociobiology. Behavioral Ecology and Sociobiology 69:1573-1574.

Tuyttens, F. A. M., S. de Graaf, J. L. T. Heerkens, L. Jacobs, E. Nalon, S. Ott, L. Stadig, E. Van Laer, and B. Ampe. 2014. Observer bias in animal behaviour research: can we believe what we score, if we score what we believe? Animal Behaviour 90:273-280.

Tuyttens, F. A. M., L. Stadig, J. L. T. Heerkens, E. Van Laer, S. Buijs, and B. Ampe. 2016. Opinion of applied ethologists on expectation bias, blinding observers and other debiasing techniques. Applied Animal Behaviour Science 181:27-33.

Tylianakis, J. M., R. K. Didham, J. Bascompte, and D. A. Wardle. 2008. Global change and species interactions in terrestrial ecosystems. Ecology Letters 11:1351-1363.

van Wilgenburg E., and M. A. Elgar. 2013. Confirmation bias in studies of nestmate recognition: a cautionary note for research into the behaviour of animals. PLoS One 8:e53548.

White, T. C. R. 1974. Hypothesis to explain outbreaks of looper caterpillars, with special reference to populations of *Selidosema suavis* in a plantation of *Pinus radiata* in New Zealand. Oecologia 16:279-301.

Zvereva, E. L., and M. V. Kozlov. 2010. Responses of terrestrial arthropods to air pollution: a meta-analysis. Environmental Science and Pollution Research 17:297-311.

Zvereva, E. L., and M. V. Kozlov. 2014. Effects of herbivory on leaf life span in woody plants: a meta-analysis. Journal of Ecology 102:873-881.

Zvereva, E. L., E. Toivonen, and M. V. Kozlov. 2008. Changes in species richness of vascular plants under the impact of air pollution: a global perspective. Global Ecology and Biogeography 17:305-319.

Zvereva, E. L., M. Roitto, and M. V. Kozlov. 2010. Growth and reproduction of vascular plants in polluted environments: a synthesis of existing knowledge. Environmental Reviews 18:355-367.

Table 1. Overview of the potential and actual effects of research methodology on the outcomes of studies of spatial patterns in insect herbivory.

Stage of research	Procedure	Expected methodological problems or predicted patterns	Compared groups of studies and the sign of the expected differences in effect sizes calculated from these studies	Conclusions of meta-analysis and relative differences in effect sizes[1]	Recommendations
Planning	Selection of plant species	Conscious and unconscious biases: selection of species that are likely to produce the expected results	Selected species > entire plant community	Confirmed: 61%	Study multiple plant species; whenever possible, obtain community-wide estimates of herbivory
	Selection of study duration	Conducting the study in a year when the	Single-year studies > multi-year studies (for studies of	Confirmed[2]: 87%	Repeat measurements of herbivory during at least

---

[1] The absolute difference in effect sizes relative to smaller effect size; reported only for significant (P<0.05) and marginally significant (0.05<P<0.08) differences.

[2] The difference is marginally significant (0.05 < P < 0.07).

		expected effect is apparent	geographic patterns)		two years
	Formulation of predictions	HARKing (Hypothesizing After Results are Known)	Studies that formulated predictions > studies that did not formulate predictions	Not confirmed	No
Sampling	Selection of individual plants	Confirmation bias	Haphazard selection > random/systematic/no selection	Not confirmed	To be on the safe side, random or systematic selection is recommended
	Selection of individual leaves	Confirmation bias	Haphazard selection > random/systematic/no selection	Confirmed: 74%	Sample all available leaves or apply strict randomization protocol
	Selection of individual leaves	Confirmation bias	Leaves collected by authors > leaves collected by technicians	Confirmed^2: 97%	Blind the hypothesis to collectors
Measurements	Assessment of leaf damage	Confirmation bias	In situ > in the laboratory	Not confirmed	No
	Assessment of	Confirmation bias	Visual > grids > image analysis	Not confirmed	No

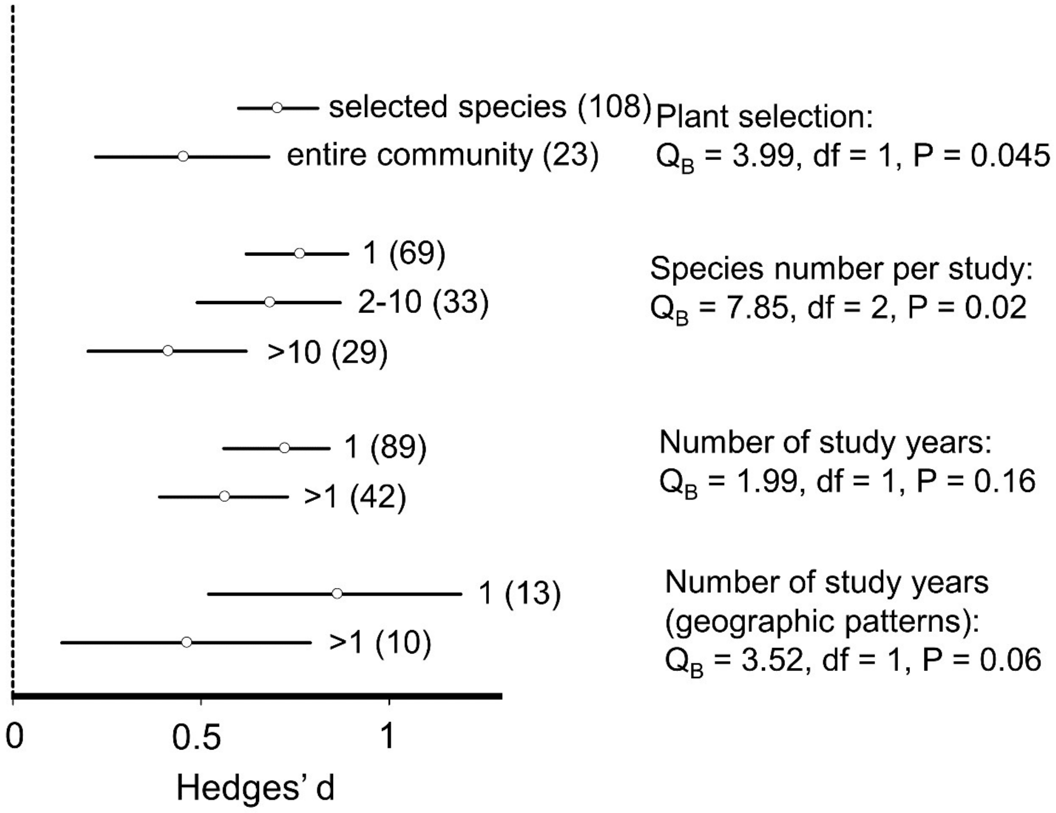
	leaf damage				
	Assessment of leaf damage	Confirmation bias	Sample origin known to measurer > sample origin unknown to measurer	Confirmed: 135%	Blind sample origin to measurers
Publication	Selection of results for publication	Publication bias	High-impact (impact factor > 2.5) journals > low-impact (impact factor < 1.6)	Opposed[2]: 58%	Increase demands for the quality and presentation of research methodology
	Peer reviewing in low-impact journals	Decrease in publication bias with time	Old publications (before 2010) > new publications (2010−2018)	Not confirmed	Increase demands for the quality and presentation of research methodology
	Peer reviewing in high-impact journals	Decrease in publication bias with time	Old publications (before 2010) > new publications (2010−2018)	Confirmed: 137%	Not needed

Figure legends

Fig. 1. Effects of the number of plant species involved in the study and the study duration on the magnitude of the detected differences in herbivory between the environments. Horizontal lines denote 95% confidence intervals; sample sizes are shown in parentheses.

Fig. 2. Effects of unconscious selection bias on differences in insect herbivory between environments. No selection: all trees per plot or all leaves per tree/branch were sampled; random or systematic selection involved pre-defined selection protocols; haphazard selection was made intuitively by a researcher without the use of any pre-defined protocol. Effects of sample collection and information about sample origin are calculated for measurements of herbivory conducted in the laboratory. In original studies, the effect size was considered negative when the outcome of the study opposed the predictions (i.e., a significant effect was observed in the direction opposite to the predicted direction). Horizontal lines denote 95% confidence intervals; sample sizes are shown in parentheses.

Fig. 3. Effects of the impact factor (IF) of the journal and of the publication year on differences in herbivory between the environments. Because we were interested in the willingness of journals to publish negative results (i.e., results opposing the predictions), the effect size in those studies was considered negative (i.e., a significant effect in the direction opposite to the predicted direction). Horizontal lines denote 95% confidence intervals; sample sizes are shown in parentheses.



selected species (108)
entire community (23)
Plant selection:
Q_B = 3.99, df = 1, P = 0.045
1 (69)
2-10 (33)
>10 (29)
Species number per study:
Q_B = 7.85, df = 2, P = 0.02
1 (89)
>1 (42)
Number of study years:
Q_B = 1.99, df = 1, P = 0.16
1 (13)
>1 (10)
Number of study years (geographic patterns):
Q_B = 3.52, df = 1, P = 0.06
0
0.5
1
Hedges' d

Fig. 1

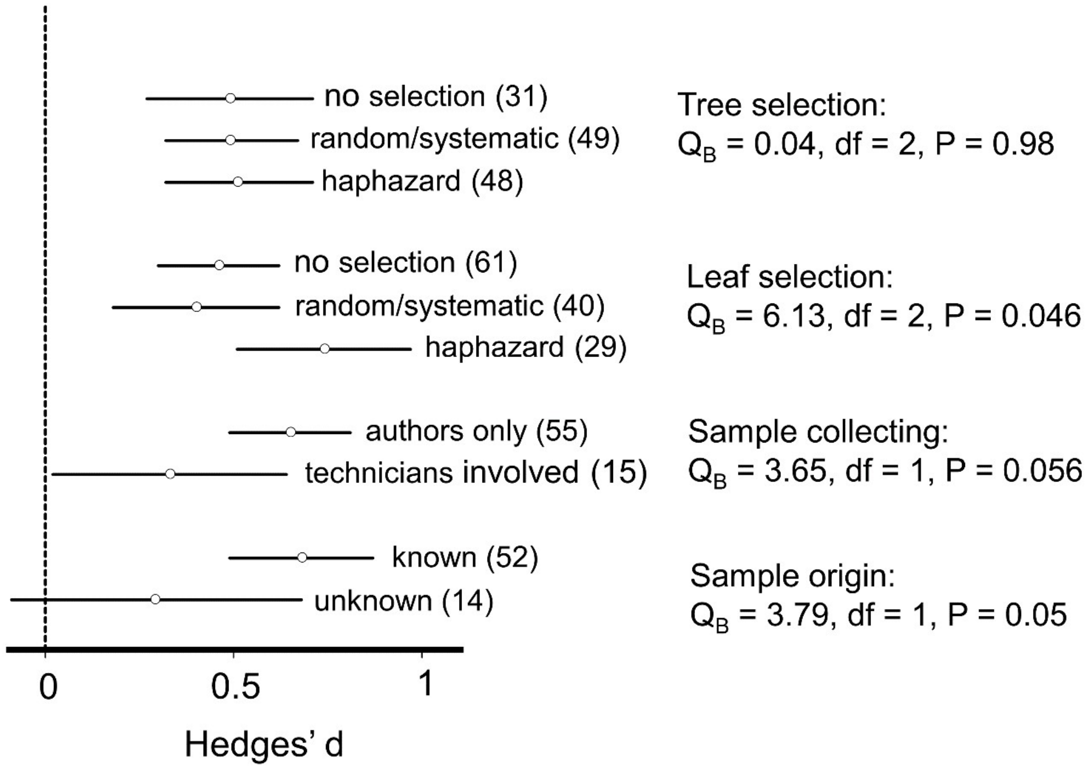

no selection (31)
random/systematic (49)
haphazard (48)
Tree selection:
Q_B = 0.04, df = 2, P = 0.98
no selection (61)
random/systematic (40)
haphazard (29)
Leaf selection:
Q_B = 6.13, df = 2, P = 0.046
authors only (55)
technicians involved (15)
Sample collecting:
Q_B = 3.65, df = 1, P = 0.056
known (52)
unknown (14)
Sample origin:
Q_B = 3.79, df = 1, P = 0.05
0
0.5
1
Hedges' d

Fig. 2

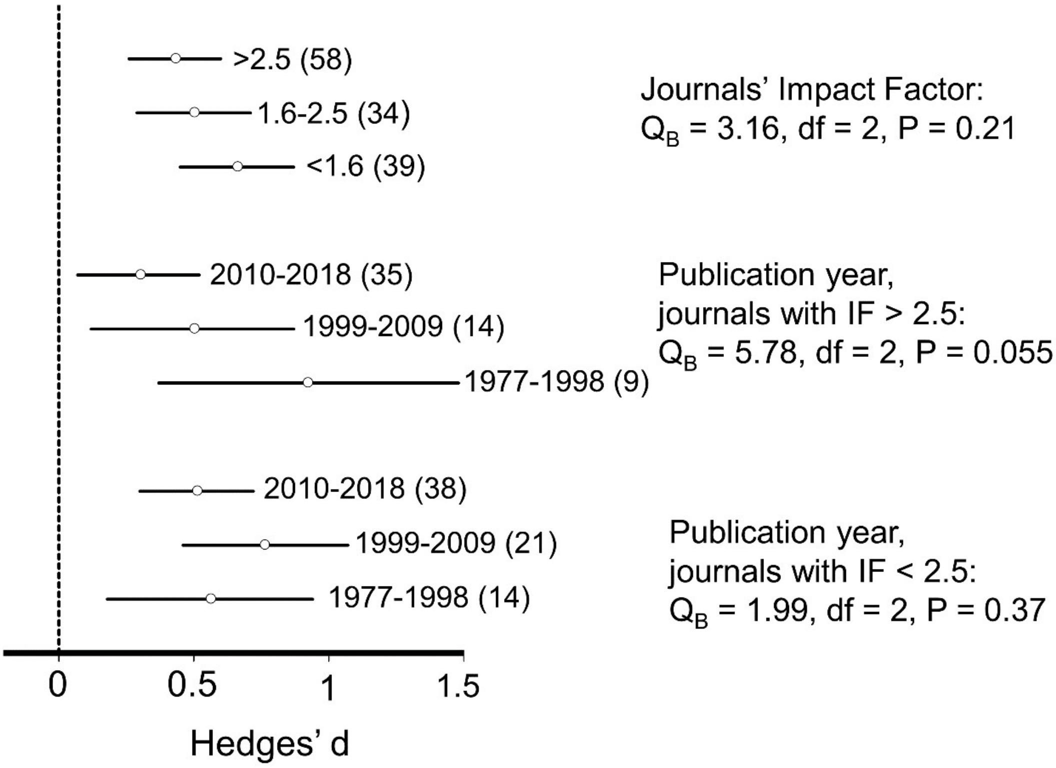

>2.5 (58)
1.6-2.5 (34)
<1.6 (39)
Journals' Impact Factor:
Q_B = 3.16, df = 2, P = 0.21
2010-2018 (35)
1999-2009 (14)
1977-1998 (9)
Publication year,
journals with IF > 2.5:
Q_B = 5.78, df = 2, P = 0.055
2010-2018 (38)
1999-2009 (21)
1977-1998 (14)
Publication year,
journals with IF < 2.5:
Q_B = 1.99, df = 2, P = 0.37
0
0.5
1
1.5
Hedges' d

Fig. 3