Meta-analysis: synthesizing research findings in ecology and evolution

Göran Arnqvist and David Wooster

The growing number of empirical studies performed in ecology and evolution creates a need for quantitative summaries of research domains to generate higher-order conclusions about general trends and patterns. Recent developments in meta-analysis (the area of statistics that is designed for summarizing and analyzing multiple independent studies) have opened up new and exciting possibilities. Unlike more traditional qualitative and narrative reviews, meta-analysis allows powerful quantitative analyses of the magnitude of effects and has a high degree of objectivity because it is based on a standardized set of statistical procedures. The first pioneering applications in ecology and evolution demonstrate that meta-analysis is both tractable and powerful.

Göran Arnqvist is at the Dept of Biology, University of New Mexico, Albuquerque, NM 87131-1091, USA, and the Dept of Animal Ecology, University of Umeå, S-901 87 Umeå, Sweden; David Wooster is at the Center for Evolutionary Ecology, University of Kentucky, Lexington, KY 40506-0225, USA.

Comparisons of sets of studies are at the heart of science: any single study is worth little if not compared and related to other similar studies. In virtually all fields of science, specific hypotheses have been addressed in multiple studies. Furthermore, studies very rarely show identical results, but instead typically differ both in the magnitude of effects and in the occurrence of significant results. Though the heuristic value of reviews and the level at which to make generalizations may be debated[1], research reviews provide the basis for conceptual syntheses and for development of general theory, and are therefore essential to scientific development. By tradition, reviews in ecology and evolution typically follow a narrative style, and the valid quantitative methods available for summaries of research domains have only recently gained attention. Narrative reviews can be seriously flawed[2,3], and understanding and integrating formal meta-analysis (MA) in our field will prove extremely important. The aim of this article is to increase the awareness and encourage the use of MA among researchers in our field (we do not provide a detailed account of the equations used to conduct an MA, which can be found elsewhere[1–5]).

What is meta-analysis?

Meta-analysis is defined as the quantitative summary of research domains, and it refers to a specific set of statistical quantitative methods that are designed to compare and synthesize the results of multiple studies[2–6]. Though some of these methods have a long history, they have been more recently incorporated into a common statistical framework[5]. Most of this development has occurred in the social sciences[4,7–9]. In the medical sciences, the number of publications using MA has increased almost exponentially in recent years[10]. Even if the current awareness of MA is very low in our field (see below), there are good reasons to expect a similar penetration into ecological and evolutionary research. In many ways, the procedures involved in MA are analogous to those of standard statistical methods, but the units of analysis are the results of independent studies rather than the independent responses of individual subjects. This has profound effects for the methods of analysis. For example, conventional statistical tests, such as standard ANOVA, regression or *t*-test, should not be applied to such data mainly because of problems with the distribution of variance (e.g. heterogeneous variances)[2]. No two studies in a set of studies are equally 'reliable'. In MA, this critical fact is accounted for by giving estimates from different studies different weights, primarily based on their sample size. Furthermore, it is important to distinguish between different quantitative review techniques; MA differs vitally from some other procedures, including 'vote-counting' and 'consistency-test' methods, that can be seriously flawed and often lead to erroneous conclusions[2,3].

Current MA offers formal methods for most types of statistical inference from a set of studies[1–5]. Meta-analysis allows the following questions to be addressed: (1) What is the combined magnitude of the effect under study? (2) Is this overall effect significantly different from zero? (3) Do any characteristics of the studies influence the magnitude of the observed effect? An MA proceeds in several steps. First, studies are gathered that address a common question or hypothesis. Second, data or test statistics from these studies are transformed into a 'common currency', called 'effect size'. Common measures of effect size are the standardized difference between means of experimental and control groups or the Pearson product moment correlation coefficient. Third, these effect sizes are combined into a common estimate of the magnitude of the effect. Fourth, the significance level of this overall effect size is computed. Fifth, the statistical homogeneity of the effect sizes is calculated. This is conducted to determine whether all studies appear to share a common effect size. Finally, the studies used in the MA can be grouped according to various characteristics of the single studies, and the effect sizes between these groups of studies can be statistically compared. A simple example of the basic procedures is given in Boxes 1–4.

Pros and cons of meta-analysis

Most researchers acknowledge the importance and usefulness of research reviews. The debate, then, is not whether such reviews are desirable, but rather which method should be used to summarize research domains. Meta-analysis has many advantages over narrative reviews. Most importantly, MA is quantitative, and thus more informative, by nature; the results are a set of numbers and probabilities that can provide reference points for development of general theory and for comparisons with other studies. Knowledge of overall effect size is also key if new studies are planned and the required number of replicates is estimated[11]. Furthermore, MA acknowledges and takes into account that all studies are not equally reliable, by quantitative weighting of studies by sample size and/or categorical measures of reliability.

Box 1. The general problem

Consider a field in ecology/evolution where 15 experimental studies (see table below) have tested for effects of a factor *x* on a response variable *y*. Only some of the studies show significant effects. Furthermore, experimental designs are not identical across studies, and different statistical methods are involved in different studies (see table). This hypothetical example is, in many ways, typical for the body of empirical work within specific research domains. The problem any reviewer of this type of data faces is to choose a method that provides the most informative and accurate summary of the results. In narrative reviews, the results below would typically be considered as 'inconsistent', 'inconclusive' or even 'conflicting'. However, the results across these different studies may actually be fully consistent, even expected, and MA provides several tractable methods for integrating the findings of these studies, which allow us actually to test for an overall effect of *x* on *y* as well as for consistency and patterns across studies (see Box 2 for a simple MA of these data).

Environment	Sample size (N)	Test statistic	Direction of effect	P-value	Estimate of statistical power	Effect size (r_p)
Terrestrial	48	$t = 2.224$	+	0.031	0.41	0.31
	66	$t = 1.661$	+	0.112	0.53	0.20
	24	$F = 5.719$	+	0.010	0.16	0.52
	26	$F = 1.543$	+	0.234	0.18	0.24
	12	$t = 0.515$	–	0.617	0.14	–0.15
	22	$t = 1.766$	+	0.092	0.22	0.36
Limnetic	8	$\chi^2 = 0.398$	–	0.528	0.13	–0.24
	24	$F = 5.879$	+	0.009	0.16	0.52
	14	$F = 1.251$	+	0.321	0.11	0.28
	72	$t = 2.071$	+	0.042	0.56	0.24
	16	$t = 0.359$	–	0.724	0.17	–0.09
Marine	14	$F = 2.046$	+	0.172	0.11	0.39
	12	$t = 0.721$	–	0.486	0.14	–0.21
	44	$\chi^2 = 7.273$	+	0.007	0.51	0.40
	28	$t = 2.570$	+	0.016	0.26	0.44

Meta-analysis is in several ways a very powerful method of analysis. In areas where effect sizes are low and/or the sample sizes within studies are restricted, MA is especially useful since it allows a highly improved control of Type II statistical errors (see Box 4). Meta-analysis is a common name for a large set of statistical procedures, which also makes it flexible. For example, it allows tests of factors that might influence or moderate the results of single studies, as well as assessments of the robustness of higher-order conclusions (see Box 3). Finally, even if a number of decisions have to be made when conducting an MA, it is less subjective than narrative reviews, since it is based on a formal, predetermined set of statistical procedures rather than individual interpretations of the data[8]. Because of the quantitative and more objective qualities of MA, it also may provide insights into research domains in which empirical data provide no clear 'consensus'. Accordingly, MA has proven to be a critical component in resolving conflicts in other fields of science[12,13].

Box 2. A meta-analysis

First, it is important to realize that we cannot expect all studies to yield significant results, even if there is a common 'true' effect. For single studies, the probability of obtaining a significant result, the statistical power, is determined primarily by the magnitude of the effect under study (effect size) and the sample size. Thus, when summarizing studies where the 'true' effect size is relatively low and/or the sample sizes are small, we would actually expect most studies to yield non-significant results (a high Type II error rate, see Box 4). This is the case in our hypothetical example, where many estimates of the statistical power of single tests, based on 'medium' effect sizes[11], show very low power (see Box 1).

In a simple MA of the data in Box 1, we first transform the outcomes of different studies to a common measure of effect size[4,29], in this case, the Pearson product moment correlation coefficient. We then (1) compute a weighted average effect size, where each study is weighted by its sample size, (2) combine the probabilities to see whether this common effect size differs from zero, and (3) assess the homogeneity of effect sizes across studies. In our case, we obtain an average weighted effect size of $r_p = 0.28$, which is highly significantly different from zero ($P < 0.001$). We then test for heterogeneity across studies by performing a diffuse test of homogeneity, but the null hypothesis that all studies share a common effect size [$\chi^2_{(14)} = 13.43$, $P > 0.5$] cannot be rejected.

A full MA would also include a set of more-focused tests of homogeneity among categorical variables, analogous to ANOVAs, such as tests for differences in effect size between studies performed in different environments or between studies using different experimental designs[2,3,6]. To conclude, this simple MA has (1) led us to conclude that there is an overall highly significant effect of *x* on *y*, (2) yielded a quantitative estimate of the magnitude of this effect, and (3) indicated that the outcome of the 15 studies are statistically indistinguishable and thus, in that sense, indeed 'consistent'.

There are also several problems and pitfalls involved with MA, most of which are shared by other types of reviews and summaries, that can potentially lead to misleading conclusions. In contrast to narrative reviews, however, MA often offers quantitative methods to address these problems. Perhaps the most universal problem is the potential bias that will result when the studies included in the MA are not representative of all studies conducted[5,14]. This may result from biases either in publication rates or in selection/retrieval of studies. Meta-analysis offers several methods for evaluating these problems and assessing the robustness of conclusions (see Box 3). If the elements synthesized in a review are not independent, this may introduce bias in the overall analysis. Such bias may be due to a number of factors, from dependence within studies when multiple tests are reported in single studies, to more diffuse reasons such as patterns of dependence among researchers, environments or taxa. Meta-analysis allows for assessment and correction for some of these problems, primarily by focused or hierarchical tests of homogeneity (analogous to ANOVAs). However, the statistical and conceptual implications of non-independence are not fully understood even in conventional statistics[15], and this is certainly also the case in MA[1–5]. Meta-analysis has also been criticized for a potential loss of information, when a research domain is summarized by a single value[1]. This criticism is not severe, however, because a proper MA not only summarizes effect size and its significance, but also offers direct statistical methods to evaluate the role of variables that may potentially influence the effect size by diffuse and focused tests of homogeneity across studies.

Meta-analysis has further been criticized for mixing 'apples and oranges'; that is, there is, as a rule, a lack of uniformity across studies. First, studies may differ in experimental condition, design or sampling unit, which may affect the results. Rosenthal[4] argued that generalizing over studies is not essentially different from generalizing over subjects within studies: mixing 'apples and oranges' is appropriate if we wish to generalize to the level of 'fruit'. More importantly, focused tests of homogeneity in an MA will, again,

Box 3. The file drawer problem

The most general problem in reviews, irrespective of which method is used, is that the studies retrieved may not be fully representative of all studies that have been carried out. This is the case if the review is based on published studies, and the probability of publication of a given study increases with increased statistical significance of its results. An extreme version of this problem, the file drawer problem, holds that the journals are filled with the 5% of the studies that show Type I statistical errors, while the file drawers in the labs contain the remaining 95% that show non-significant results.

One of the strengths of MA is that it offers methods for addressing this problem, in contrast to non-quantitative reviews[30–33]. For example, because an MA yields a quantitative estimate of the effect size and its level of significance, it allows us to calculate the number of studies showing an average zero effect size that must be in the file drawers before the overall probability is brought below $P=0.05$ (Refs 2,4). This number of studies represents the review's tolerance for null results, and the conclusions of the review can then be assessed by evaluating whether the tolerance level, in relation to the number of studies on which the review is based, is small enough to constitute a file drawer 'threat'. The tolerance level can be assessed either by comparing it to recommended tolerance limits[2,4], or by comparing it to surveys of the actual number of studies in the file drawers of researchers within the domain. In our hypothetical case (see Boxes 1 and 2), 121 zero-result studies must be hidden in file drawers, which exceeds recommended tolerance level limits, and we may thus regard our conclusion of an effect of *x* on *y* as robust to the file drawer problem.

Box 4. Meta-analysis and Type II error rates

The primary concern in conventional statistics is to avoid cases where true null hypotheses are rejected (Type I errors). Consequently, most procedures for statistical inference have been more or less designed to control for Type I error rates. In contrast, dealing with cases where we fail to reject false null hypotheses (Type II errors) is much more problematic. This poses a huge potential problem, and there is little doubt that Type II errors are far more frequent than Type I errors, not only in our domain[34], but in most areas of science[2,11]. A major step towards understanding the importance of Type II errors is consideration of the statistical power of single tests (see Box 2), but the methods for actually controlling for Type II errors are still crude. This fact is especially troublesome in areas where failure to reject false null hypotheses may have large and important impacts, such as in conservation biology and medicine, where Type II errors could lead to usage of detrimental exploitation policies of natural resources or rejection of effective medical treatments[21,34].

A major beneficial feature of MA is that it offers a unique and highly improved control of Type II error rates[2,4,5]. Even if the number of studies in the MA is modest, Type II error rate is drastically reduced. As an illustration, an overall effect size may be significant in an MA of a set of consistent studies, even if none of the single studies shows significant results (i.e. all show Type II errors)[29]. In our hypothetical example (Boxes 1 and 2), we found an overall effect of *x* on *y* that was consistent across studies. Since nine of the 15 studies showed non-significant results, our MA indicates a Type II error rate of roughly 60% among single studies. Because of its ability to deal with Type II error rates, MA is especially attractive when the Type II errors are either detrimental (see above) or frequent (low statistical power of single studies for some reason).

enable assessments and tests of differences between subgroups of studies[2,6,16,17]. Second, MA has been criticized for mixing 'good' and 'bad' studies. But MA actually offers formal methods both for defining reliability of single studies and for accounting for differences in reliability across studies. Because of its ability to control for Type II error rates, MA has also been criticized for exaggerating the significance level in a combined analysis. Even though a decrease in statistical error rate cannot be viewed as a legitimate criticism (Box 4), various methods for combining probabilities are available and should nevertheless be chosen carefully, since any given method may be more or less appropriate in special cases[4]. It is crucial to realize that MA has important conceptual limitations. It has been suggested that MA may actually offer an alternative to complex experimental designs, and that a summary of several simple experiments may be more informative than conducting a few complex ones[16]. However, while MA may be complementary to multifactorial experiments, it cannot replace such studies: it is not, for instance, possible to disentangle the separate effects of two correlated independent variables (e.g. age and size) by an MA of multiple studies that each deal with either of these two variables.

In conclusion, MA offers several profound advantages over narrative reviews. Furthermore, most of the potentially serious problems involved in MA are shared with other forms of reviews, but can be at least partly ameliorated in MA by (1) detailed consideration of how studies are selected/retrieved for the MA, (2) tests of robustness of conclusions (see Box 3), (3) careful consideration of potential interdependences across studies, and (4) assessments of differences between subgroups of studies by focused tests of homogeneity.

Meta-analysis in ecology and evolution

To date, MA has been used in only a handful of reviews in ecology and evolution (see Box 5). Below, we briefly review three of these papers, to illustrate the range of topics and questions amenable to MA and the types of results that an MA provides. These papers follow a common general outline: identification of a critical hypothesis or controversy, determination of the effect size of the factor in question, and examination of the role that various ecological and methodological factors play in explaining variability in effect sizes.

Competition is a key concept in population and community ecology. Two narrative reviews of competition concluded that it occurs 'frequently' in natural systems[18,19]. Neither review was able to make statements about the magnitude of the effect that competition has on organisms. Thus, the question remained, to what extent does competition affect density/biomass of competing organisms? Gurevitch *et al.*[16] performed an extensive MA of field studies on competition, and were able to show that competition had a strong and highly significant impact on plant and animal biomass across all comparisons. The studies used in this review were partitioned into subgroups, to enable assessments of the role of various ecological and methodological factors. For example, effect sizes were compared across trophic levels, revealing the effect of competition among herbivores to be stronger and more variable compared to primary producers and carnivores. This pattern actually conflicts with standard theory, which predicts that carnivores and primary producers should show strong effects of competition, whereas herbivores should not because of predator limitation[20]. Gurevitch *et al.* also showed an impact of the experimental design on the observed effect sizes. Studies of high reliability (i.e. long-term, well-planned studies, large sample size) showed a lower variability in effect size compared to less-reliable studies, in accordance with MA expectations.

Meta-analysis has also been used to test classic theories in ecology/evolution. David Lack's clutch size hypothesis states that animals should tend to produce optimal clutches, defined as the clutch size that produces the highest number of surviving offspring[17]. Empirical results of studies that manipulate clutch size in birds have produced variable results: some studies show that birds are capable of rearing additional offspring, while others suggest that they are not. A narrative review would most likely conclude that because of the inconsistent results of experiments, no general conclusions can be drawn. Using MA, VanderWerf[17] compared the standardized results of clutch-size manipulation experiments, and contrary to Lack's hypothesis found that, in general, birds are capable of raising enlarged broods. Further analyses revealed two key features of offspring production in birds. First, studies conducted over more than one nesting season were less likely to show that adults were capable of raising enlarged clutches, indicating that birds might optimize clutch size over periods greater than one year. Second, altricial birds were less likely to raise enlarged clutches than precocial birds, suggesting that altricial young place greater demands on parents than precocial young.

Box 5. Meta-analysis in ecology and evolution

Although only a handful of reviews in ecology and evolution have used MA to date, those that have address a wide variety of questions. Furthermore, the use of MA in our field seems to be rapidly increasing: the publication rate so far is one published paper in 1991, two in 1992 and four in 1994. To illustrate the conceptual range of these reviews, we provide their major question asked and the subsequent results (see below). Effect sizes are the standardized difference between control and treatment means (d). As benchmarks, effect sizes of 0.80 are considered as strong effects, 0.50 as moderate effects, and 0.20 as weak effects[11]. All effect sizes given are significant ($P \leq 0.05$).

Question	Conclusion
Does competition affect biomass?	Yes, a strong but variable effect[16] ($d = 0.80$).
Do predators influence the density of benthic prey in stream communities?	Yes, a weak to moderate effect[35] ($d = 0.39$).
Does age influence laying date and clutch size in great tits and pied flycatchers?	Yes, young females lay later ($d = 0.30$ and $d = 0.45$) and have smaller clutches[36] ($d = 0.31$ and $d = 0.81$).
Does Lack's clutch size hypothesis hold for birds?	No, birds are capable of raising enlarged broods[17] ($d = 0.55$).
Do helminth parasites induce changes in host behavior?	Yes, parasites have moderate effects on host activity and microhabitat choice[37] ($d = 0.50$ and $d = 0.59$).
Does crop diversification cause a decline in the population density of insect pests?	Yes, a moderate effect[22] ($d = 0.27–0.50$).
Does selective logging decrease the density of breeding birds?	Yes, undisturbed forests have higher densities[21] ($d = 0.59$).

Meta-analysis has a large potential value in applied ecology as well[21,22]. A potential method of insect pest control in agriculture is crop diversification, which should lead to lower densities of insect pests that specialize on any single plant species. Empirical tests of the impact that crop diversification has on insect pest populations have shown variable results: some studies suggest that crop diversification does lead to smaller insect populations, while others have failed to detect such an effect. Tonhasca and Byrne[22] used MA to assess the value of crop diversification as a means of biological control. The analysis revealed that crop diversification indeed leads to a moderate but significant decline in insect pest densities. This result suggests that while crop diversification helps limit crop damage, its effect alone may be too weak, and additional methods may be needed to achieve the desired effects.

These studies all illustrate three important benefits provided by MA: their conclusions were reached by applying predetermined statistical procedures, they provide quantitative measures of the magnitude of effects, and they were able to test for effects of ecological and methodological factors by analyzing subgroups of studies.

The potential for meta-analysis in ecology and evolution

Meta-analysis will prove most useful in areas of ecology and evolution where (1) there is a moderate to large amount of empirical work available, (2) the results are variable across studies, (3) the expected magnitude of the effect is relatively weak, and/or (4) the sample sizes of individual studies are limited for some reason. Many ecologists and evolutionary biologists most likely will recognize subjects in their own research area in which these circumstances apply. Since MA is also fairly easy to perform, we anticipate that MA will rapidly become a common and important method of summarizing research in ecology and evolution[1]. Below, we suggest some broad areas in which MA might be applied successfully, to illustrate the variety of areas in which MA would prove helpful.

Behavioral ecology is an area rich in theory and empirical studies, and many areas in behavioral ecology qualify for MA. For example, optimal foraging theory has received a lot of theoretical and empirical attention. It should be possible to use MA to determine whether, across numerous studies, organisms seem to conform to predictions made by optimality theory (e.g. rate maximizing, switching). Meta-analysis might also be useful to determine how social behavior influences foraging rates or reproductive success. For example, is dominance rank generally correlated with foraging and/or number of matings? Such a correlation has been found in primates[23], but it would be enlightening to expand this analysis to include other taxa.

Community ecology is an area of research in which MA has already been successfully applied to examine the impact of competition on biomass[1,16]. We believe that there are many more areas in community ecology that would equally benefit from MA. For example, there are many interactions between species that theoretically, and in some cases empirically, have important impacts on community structure (e.g. higher order interactions, indirect effects, parasitism, mutualism). MA could allow determinations of the strengths of these interactions and how they vary under different ecological and methodological conditions.

Meta-analysis should also prove useful in a variety of subdisciplines of evolutionary biology, including life history theory. General life history theory predicts negative covariances between life history characters, owing to evolutionary limitations and trade-offs. Empirical evaluations of whether or not such trade-offs exist are often considered to be 'inconsistent'. For example, some field and laboratory evidence suggests a negative correlation between reproduction and survival, while other studies suggest a positive correlation, and a third group fail to find any correlation[24]. Here, MA would certainly help in testing for general effects as well as for differences between taxa or methods. Also, MA could be used to compare the degree of additive genetic variation among traits of various types. Traits closely related to fitness are expected to show less variability, but empirical results are inconsistent[25,26]. Furthermore, several areas of sexual selection should be suitable for MA. For example, MA could enable detection and analysis of weak patterns of nonrandom mating across multiple studies. A future MA could also help resolve the debate over whether female mate preferences and male sexual ornamentations are genetically correlated[27]. Another dispute that may be suitable for MA is whether females in general choose males with certain phenotypic characteristics (e.g. low parasite loads or low levels of fluctuating asymmetry) as mates[28].

Conclusions

Reviews are at the heart of scientific development: they provide general conclusions and guidance for future research. Meta-analysis is in many ways a superior method of comparing and summarizing the results of multiple studies.

There are, however, several potential pitfalls involved when performing an MA, most of which are shared with other review methods, and great care should be taken to avoid and assess these problems. Meta-analysis is still rare in our domain, but the first applications show that it can successfully help address a variety of questions. Because the empirical basis is now sufficient for many topics and the methods of MA are widely available, the stage is set for a variety of thorough and informative reviews in which major questions in ecology and evolution are addressed.

Acknowledgements

We thank B. Brodie, J. Brown, P. Crowley, J. Gurevitch, F. Johansson, A. Kodric-Brown, A. Sih, R. Thornhill, J. Wagner and P. Watson for discussions and constructive comments on previous drafts of this paper. The support to G.A., from The Swedish Natural Science Research Council, The Crafoord Foundation (The Royal Swedish Academy of Sciences), The Fulbright Commission and The Swedish Institute, and to D.W., from The Graduate School, The University of Kentucky, USA, is gratefully acknowledged.

References

1 Gurevitch, J. and Hedges, L.V. (1993) in *Design and Analysis of Ecological Experiments* (Scheiner, S.M. and Gurevitch, J., eds), pp. 378–398, Chapman & Hall
2 Hunter, J.E. and Schmidt, F.L. (1990) *Methods of Meta-Analysis*, Sage Publications
3 Hedges, L.V. and Olkin, I. (1985) *Statistical Methods for Meta-Analysis*, Academic Press
4 Rosenthal, R. (1991) *Meta-Analytic Procedures for Social Research*, Sage Publications
5 Cooper, H.M. and Hedges, L.V., eds (1993) *Handbook of Research Synthesis*, Russel Sage Foundation
6 Mullen, B. (1989) *Advanced BASIC Meta-Analysis*, Lawrence Erlbaum
7 Glass, G.V., McGraw, B. and Smith, M.L. (1981) *Meta-Analysis in Social Research*, Sage Publications
8 Mullen, B. and Rosenthal, R. (1985) *BASIC Meta-Analysis: Procedures and Programs*, Lawrence Erlbaum
9 Rosenthal, R. and Rosnow, R.L. (1991) *Essentials of Behavioural Research: Methods and Data Analysis*, McGraw-Hill
10 Selden, C. (1992) *Curr. Bibliogr. Med.* 1992/13
11 Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum
12 Mann, C.C. (1990) *Science* 249, 476–480
13 Mann, C.C. (1994) *Science* 266, 960–962
14 Wachter, K.W. and Straf, M.L., eds (1990) *The Future of Meta-Analysis*, Russel Sage Foundation
15 Hurlbert, S.H. (1984) *Ecol. Monogr.* 54, 187–211
16 Gurevitch, J., Morrow, L.L., Wallace, A. and Walsh, J.S. (1992) *Am. Nat.* 140, 539–572
17 VanderWerf, E. (1992) *Ecology* 73, 1699–1705
18 Connell, J.H. (1983) *Am. Nat.* 122, 661–696
19 Schoener, T.W. (1983) *Am. Nat.* 122, 240–285
20 Hairston, N.G., Smith, F.E. and Slobodkin, L.B. (1960) *Am. Nat.* 94, 421–425
21 Fernandez-Duque, E. and Valeggia, C. (1994) *Conserv. Biol.* 8, 555–561
22 Tonhasca, A. and Byrne, D.N. (1994) *Ecol. Entomol.* 19, 239–244
23 Cowlishaw, G. and Dunbar, R.I.M. (1991) *Anim. Behav.* 41, 1045–1056
24 Roff, D.A. (1992) *The Evolution of Life Histories: Theory and Analysis*, Chapman & Hall
25 Roff, D.A. and Mosseau, T.A. (1987) *Heredity* 58, 103–118
26 Houle, D. (1992) *Genetics* 130, 195–204
27 Breden, F., Gerhardt, H.C. and Butlin, R.K. (1994) *Trends Ecol. Evol.* 9, 343
28 Andersen, M. (1994) *Sexual Selection*, Princeton University Press
29 Kirby, K.N. (1993) *Advanced Data Analysis with SYSTAT*, Van Nostrand Reinhold
30 Simes, R.J. (1987) *Stat. Med.* 6, 11–29
31 Iyengar, S. and Greenhouse, J.B. (1988) *Stat. Sci.* 3, 109–135
32 Dear, K.G.B. and Begg, C.B. (1992) *Stat. Sci.* 7, 237–245
33 Hedges, L.V. (1992) *Stat. Sci.* 7, 246–255
34 Shrader-Frechette, K.S. and McCoy, E.D. (1992) *Trends Ecol. Evol.* 7, 96–99
35 Wooster, D. (1994) *Oecologia* 99, 7–15
36 Järvinen, A. (1991) *Ibis* 133, 62–67
37 Poulin, R. (1994) *Anim. Behav.* 48, 137–146

Trends in Ecology & Evolution editorial policy

Trends in Ecology & Evolution (*TREE*) is a current-awareness magazine publishing news, reviews and comment on research activity at the front of ecology and evolutionary biology. It is not a journal for the publication of original results, hypotheses, syntheses or meta-analyses. *TREE*'s authors are encouraged to perceive their role as reporters, chroniclers and commentators, rather than using *TREE* as a vehicle to publish their own work in a different format.

News & Comment articles report on specific new developments published in the primary literature; they also cover conferences. The former, but not the latter, are usually peer reviewed. **Reviews** chronicle recent progress in particular areas of research. **Reviews** and **Perspectives** essays are always peer reviewed.

Articles in *TREE* are commissioned by the Editor. Manuscripts may nevertheless be rejected by the peer and editorial review process: commissioning does not guarantee publication.