

Dynamic Ecology

Multa novit vulpes

Why do so many ecologists overestimate how informative small meta-analyses are about the mean effect size?

Posted on **February 10, 2021** by **Jeremy Fox**

Recently, I showed poll results and other data tentatively making three points:

1. The typical ecological meta-analysis includes 50ish effect sizes
2. Many ecologists think that an ecological meta-analysis only needs 50ish effect sizes in order to provide a stable estimate of the mean effect size
3. In fact, ecological meta-analyses typically need 250-500 effect sizes, or maybe even more, to provide a stable estimate of the mean effect size

In other words, many ecologists overestimate how informative most meta-analyses are about the mean effect size. Why?

I don't think it's because most ecologists haven't ever thought about this issue (is it?). I don't think most ecologists just subconsciously assume that however big a typical ecological meta-analysis is, that's surely big enough (is it?). And I don't think it's because most ecologists are unaware of the assumptions and statistical properties of the hierarchical random effects models used for most ecological meta-analyses (is it?). So what's the thinking? Here are the lines of thinking I've heard, in our comment threads and in correspondence with colleagues, and my responses to them.

tl;dr: I don't come to bury Caesar—meta-analyses. Meta-analyses can be useful! They're a good tool for ecologists to have in our toolbox! I'm not one of those rare-but-annoying ecologists who think that meta-analyses and the ecologists who do them are somehow Bad. I just think that, for certain purposes, meta-analyses in ecology aren't as useful as many ecologists think they are. I think it would be healthy to have a clear-eyed discussion about that. Maybe there are ways to do meta-analyses better. Or better ways to accomplish some of the goals we're trying to accomplish with meta-analyses.

1. **Small meta-analyses aren't all that informative in themselves, but they're useful because they identify topics on which further research is needed, and motivate ecologists to conduct further research on those topics.** To which, sure, could be, in some cases. I don't know if that's typically the case, though. Personally, I don't think ecological meta-analyses tend to be a very good source of [stylized facts](#). And anecdotally, I don't think publication of a small meta-analysis often tends to inspire others to collect additional effect size estimates that could be incorporated into some future, larger meta-analysis on the same topic. Ecologists' choices of research topic depend on lots of factors; I doubt that publication of a small meta-analysis does much to shape ecologists' choices of research topic. And even if some ecologists do respond to a small meta-analysis by going out and collecting further effect sizes, those further effect sizes rarely end up in an updated meta-analysis. Only a very small fraction of meta-analyses in ecology are updates of a previous, smaller meta-analysis on the same topic.
2. **Small meta-analyses are useful when they're focused: when they only include effect sizes from very similar studies. Small, focused meta-analysis provide precise estimates of mean effect size under specific conditions. Too many ecological meta-analyses are over-broad. Averaging together a heterogeneous mix of studies in different study systems using different methods amounts to comparing apples and oranges.** This is a very intuitively-appealing thought—it appeals to my intuitions too! But those intuitions are wrong. I don't think ecologists can actually tell which studies are truly "similar" to one another (i.e. low "heterogeneity"), so that a meta-analysis of those studies will yield a precise, stable estimate of the mean effect size despite including only a few studies. I say that for four reasons. First, there are few ecological meta-analyses—not even small, purportedly "focused" ones—for which most variation in effect size is attributable to sampling error rather than to heterogeneity. And those few low-heterogeneity meta-analyses aren't more "focused" than other meta-analyses in any obvious way. Second, [within-study heterogeneity in effect size is about as large on average as among-study heterogeneity](#). That is, in a hierarchical random effects meta-analysis, there's often as much, or even more, heterogeneity among effect sizes *from the same paper* as there is among effect sizes from different papers. Even though different effect size estimates from the same paper usually were collected by the same investigators, using the same methods, in the same or nearby locations, etc. Third, small ecological meta-analyses do not typically report more precise estimates of mean effect size than larger meta-analyses do. In fact, just the opposite. So even if you are reducing heterogeneity by conducting a small meta-analysis that only includes very similar studies, you're likely just trading heterogeneity for sampling error. Fourth, a personal anecdote. Many years ago, I did a laboratory microcosm experiment on trophic cascades ([Fox 2007](#)). It involved protist predators, their protist prey, and bacteria as the basal trophic level. The species were purchased from biological supply companies and so lacked any coevolutionary history, they were growing under constant environmental conditions in an artificial environment without any other species present, etc. Quite different in all sorts of ways from any natural system! But yet, when I calculated mean effect sizes to measure the strength of trophic cascades in my experiment, the effect sizes were almost bang-on identical to the mean strength of trophic cascades in field studies, according to the meta-analysis of Borer et al. 2005. Now, I don't think this shows that my microcosms actually were "realistic". Rather, it shows that no ecologist—including me!—knows enough about the

Why do so many ecologists overestimate how informative small meta-analyses are about the mean effect size? | Dynamic Ecology determinants of trophic cascade strength to confidently guess which studies of trophic cascades will produce similar effect sizes on average. For those four reasons, I've decided that ecologists can't actually tell which studies are sufficiently "similar" to one another to belong in the same, narrowly-focused meta-analysis. I don't think ecologists can tell when they're comparing "apples to apples" rather than "apples to oranges". Because different ecological "apples" are different from one another in all sorts of ways, most of which ecologists are unaware of. And because ecologists can't really tell the difference between apples, oranges, pictures of apples, and jars of applesauce.*

3. **The grand mean effect size usually isn't of scientific interest, because the mean effect size will depend on the values of one or more moderator variables. If the grand mean effect size estimate is noisy, that's a sign that you need to add moderator variables to the meta-analysis, so as to explain most of the noise.** This point is closely related to #2, obviously. And so my response to it is along the lines of my response to #2. Now, I haven't compiled data on moderator variables in ecological meta-analyses. But speaking as someone who's skimmed a lot of ecological meta-analyses: moderator variables in ecological meta-analyses hardly ever seem to explain more than a tiny fraction of variation in effect size, even when that fraction is significantly >0. I can't recall ever seeing an ecological meta-analysis in which conditioning on moderator variables really did all *that* much to improve precision of estimated mean effect sizes. If you think I'm forgetting some examples, by all means name them! I'd be super-interested to be pointed towards ecological meta-analyses in which moderator variables explain a big chunk of the variation in effect size! But as best I can tell based on the meta-analyses I've seen, moderator variables in meta-analyses are rather like genetic loci of small effect. Genetic variation in highly-heritable quantitative traits (e.g., human height) is underpinned by genetic variation at many loci. There may be a few variants of "large" effect, but even the largest explains only (say) 10% of the heritable variation in the quantitative trait of interest. Most of the genetic variants at most of those loci are so rare and/or have such small effects that we have little hope of identifying them, even in genome-wide association studies with massive sample sizes ([Young 2019](#), [Pallares 2020](#)). This is the so-called "missing heritability" problem: we know that quantitative traits like human height are heritable, but we can't identify the genetic variants that, collectively, underpin most of that heritability. That's why quantitative genetics (still) exists: rather than trying to identify the genetic variants that matter, just *statistically summarize* genetic variation with a genetic variance and be done with it. I increasingly wonder if ecological meta-analysts ought to go the same route as quantitative genetics. Just give up on trying to explain variation in effect size with moderator variables.
4. **Maybe ecological meta-analyses used to be small and noisy. But these days, they include many more effect sizes than they used to. Further, many of those effect sizes have less sampling variance than they used to, because primary studies these days use larger sample sizes and more rigorous methods than they used to. Right?** Wrong. (Sorry!) Over time, the number of effect sizes per ecological meta-analysis has become more *variable*. Yes, there are some meta-analyses published these days that are much bigger than any meta-analysis published back in the 1990s or oughts. But there are also some meta-analyses published these days that are much *smaller* than any meta-analysis published back in the 1990s or oughts. And no, effect sizes published these days do not have smaller

Why do so many ecologists overestimate how informative small meta-analyses are about the mean effect size? | Dynamic Ecology sampling variances on average than those published decades ago. Or if they do, the trend of declining sampling variances over time is very weak and noisy.

5. **The most important things are to be able to reliably estimate the sign of the mean effect size, and infer that the mean isn't truly zero. As long as we can infer with reasonable reliability that the mean effect size is >0 or <0 , we don't need to worry too much if our estimate of the absolute magnitude of the mean is imprecise or unstable.** Sure, fine, I guess. Though this strikes me as a rather unambitious point of view. The other difficulty for this point of view is that, even in cases where we're pretty confident in the estimated sign of the mean effect size, the distribution of observed effect sizes often will include many effect sizes of each sign. Do we care about the sign of the mean effect size? Or about how common positive effect sizes are relative to negative effect sizes? Because those are two different questions.
6. **Even though we can't precisely estimate the grand mean effect size, or explain much of the variation in effect size with moderator variables, the small fraction of variation in effect size that we can explain with moderator variables often is of substantial scientific interest. For instance because we want to test hypotheses or predictions about how effect size depends on some particular moderator variable.** Sure, absolutely. I completely agree that it's often of scientific interest to test hypotheses about the effect of variable X on variable Y, even if variable X explains only a small fraction of the variation in variable Y. Most of my own research is those sorts of studies! It's just important to be clear in your own mind, and make clear to your readers, that that's the sort of study you're doing. Don't say or imply that X explains lots of the variation in Y if in fact it doesn't.
7. **Yes, our estimates of mean effect size are super-variable, and no, we'll never explain much of the variation in effect size (well, maybe unless we switch to different statistical methods), but that's what makes ecology fun! If ecology weren't full of difficult-to-explain variation, it'd be physics, and who wants to be a physicist?** This point of view is one Brian discussed in [his old post on "multicausality"](#), so I'll just point you to Brian's old post rather than trying to summarize it. It's a good discussion, you should go read it!

So, what do you think? Is there a problem here with how many ecologists use and interpret meta-analyses? If so, what should (or could) we do to improve matters? Looking forward to your comments.

*That high-pitched sound you hear is the “apples to apples” analogy screaming in pain as I torture it mercilessly.

This entry was posted in [Uncategorized](#) by [Jeremy Fox](#). Bookmark the [permalink](#) [<https://dynamicecology.wordpress.com/2021/02/10/why-do-so-many-ecologists-overestimate-how-informative-small-meta-analyses-are-about-the-mean-effect-size/>].



About Jeremy Fox

I'm an ecologist at the University of Calgary. I study population and community dynamics, using mathematical models and experiments.

[View all posts by Jeremy Fox →](#)

25 THOUGHTS ON "WHY DO SO MANY ECOLOGISTS OVERESTIMATE HOW INFORMATIVE SMALL META-ANALYSES ARE ABOUT THE MEAN EFFECT SIZE?"



Mark Vellend

on February 10, 2021 at 1:36 pm said:

Sorry if this was addressed in one of the previous posts, but...what if we step back from the question of a stabilized mean effect size specifically and ask: "is it worth doing a meta-analysis with 50 effect sizes?". If there were only 10, the Intro section of a new paper can provide a good overview of what is known on the topic via qualitative description. If there are 50, that's not possible, and it sure seems useful to have a quantitative analysis of the current state of knowledge when 50 studies have been done. Especially if one is doing truly applied research where decisions are being made based (in part) on data (e.g., effects of Covid vaccines). Your analysis shows that we shouldn't over-interpret a mean = 0.4, but I can well imagine a confidence interval of 0.1-0.7 being quite useful, and the formal analysis is also likely to reveal what kinds of systems are underrepresented so far (the latter doesn't require the formal meta-analysis, but a useful biproduct all the same). So, what's your answer to the question: is it worth doing a meta-analysis with 50 effect sizes?



Jeremy Fox

on February 10, 2021 at 2:11 pm said:

"Is it really worth doing a meta-analysis with (say) 50 or fewer effect sizes?" is definitely a question I think we should be asking ourselves!

I agree with you that there's no point to very small meta-analyses—say, 10ish effect sizes or less. Such meta-analyses aren't that rare these days; they're often reported in papers that report several meta-analyses. Someone decides to (say) do separate meta-analyses for effects of variable X on species richness, species evenness, and total abundance, and there are only 10 studies of total abundance or something. In such cases, I'd prefer meta-analysis authors to just say "we looked into doing a meta-analysis on total abundance as well, but there were only 10 studies, so we didn't bother."

Is it worth doing a meta-analysis with, say, 50 effect sizes rather than 10? [thinks] You know what, I'm going to grasp the nettle here and say no. Or not usually. I can imagine a few specific circumstances in which a 50-effect meta-analysis might be useful. You mention the main one: applied contexts where the meta-analysis directly addresses a specific management question. Without wanting to downplay the importance of those specific contexts, I don't think they're that common. Meaning, most ecological meta-analyses that only include 50ish effect sizes do not directly address some specific management question. And I'm not convinced that most of them have some other strong rationale for existing.

Which is not entirely or even mostly meta-analysts' fault! I mean, in part it is—meta-analysts have a wide choice of exactly what question to ask, and so have a wide choice of exactly what studies to include. But it's mostly down to the fact that there aren't **that** many questions ecologists want to ask—even very broadly-framed questions—for which one can extract, say, 250+ effect sizes from the ecological literature. Because if there were many such questions that had been **that** well-studied, meta-analysts would've asked them already, and the typical ecological meta-analysis would have 250+ effect sizes! Ecologists' collective research efforts just haven't been **that** tightly focused. We ecologists mostly do our own thing, that's only loosely connected to the things most other ecologists are doing.

One possible response to this is to have much more central coordination of ecological research effort. Funding agencies could for instance decide to put all their money into what would basically be “contract research”. Fund just a few very specific, tightly-focused, centrally-coordinated research programs, that will have massive sample sizes. A few ecological equivalents of the Manhattan Project, or the Large Hadron Collider to look for the Higgs boson. Get a few precise estimates of a few mean effect sizes, rather than many imprecise estimates of many mean effect sizes. I don't think I'd be in favor of that, it seems worse than the status quo to me. But maybe that's just a failure of imagination on my part. I find it hard to imagine how the field of ecology would operate in that scenario. What would the focal questions and research projects be? How would they be chosen? I can't even wrap my mind around it.

Another possible response is “do more cheap distributed experiments”. More NutNets. Everybody **mostly** still does their own thing—but they also spend part of their time contributing to a big, focused, centrally-coordinated thing with a massive sample size. Ecology's moving in that direction, I think. I think it's still an open (and interesting!) question how much that approach cuts down on heterogeneity of effect size estimates (by forcing everyone to use exactly the same methods and collect exactly the same data at every study site).

Another answer is the one I'd give: there's more to “generalizing” or “synthesizing” in ecology than just “meta-analysis”. So if a typical ecological meta-analysis is too small for us to really learn all that much from it, well, maybe that's a reason not to focus so much on meta-analyses. Don't abandon meta-analyses entirely, just give more of our collective effort and attention to other ways of attaining “generality” or “synthesis”. Shameless self-promotion:

The five roads to generality in ecology (UPDATED)

Ecology is infamously variable. Every individual organism, site, time, species, population, community respects. So how do you achieve generality? How do you discover (or impose!) some sort of unit



Dynamic Ecology

Poll results: the many ways ecologists seek generality (and why)

Ecologists study lots of different things, and those things vary a lot in all sorts of ways. No two sites/webs/whatever are exactly alike. Should ecologists seek generalities about the diverse, heterogeneous world?



Dynamic Ecology

Although one can of course question how useful some of our other ways of achieving “generality” or “synthesis” are. I confess I’m not a fan of the sorts of opinion and perspectives paper that’s based around a box-and-arrow diagram, showing that various ecological concepts might affect or relate to various other ecological concepts. Those seem like advertorials for the worldviews or research programs of the authors, rather than scientific advances. If the alternative to “publish lots of small meta-analyses” was “publish lots of opinion/perspective pieces based on box-and-arrow diagrams”, I’d feel pretty unhappy with those choices! Fortunately, those obviously aren’t the only two choices.



Tim Parker

on [February 10, 2021 at 2:53 pm](#) said:

I find Mark's argument that, with 50 studies, meta-analysis serves an important role in summarizing results. Quantitative review may be imperfect for the reasons described above (and others as well), but qualitative reviews in the absence of quantitative summaries seem even less reliable to me. A quantitative summary could be as simple as a forest plot with effects plotted from all studies (though I am happy to admit that I would like to see some means as well, even if they are imprecise). Actually, I think it would be really useful if every time we published an empirical study, we included a forest plot with all the other published effect sizes from 'similar' studies (however we define them) to put our work in context and to update the state of accumulated knowledge. Making such a forest plot is time consuming with a large body of literature that has not already been summarized, but if a meta-analysis has been conducted, or if the summarizing is iterative (it becomes a norm in the field), then this would be fairly easy.

I have more to say on the topic of this post, including both points of agreement and disagreement, but no time at the moment. Maybe later.

**Jeremy Fox**on **February 10, 2021 at 3:05 pm** said:

"qualitative reviews in the absence of quantitative summaries seem even less reliable to me."

Oh, I agree that a qualitative review *that tries to infer the same things that a meta-analysis tries to infer* isn't an improvement on a meta-analysis.
Probably it's worse in many cases.

Interesting idea for putting new studies in context graphically. Will need to think more about that.

**Jeremy Fox**on **February 10, 2021 at 2:20 pm** said:

I was amused just now to think back to 2011 me—me back when I first started blogging. Back then, I definitely would've titled this post something deliberately provocative, like "Most ecological meta-analyses are too small to be worth publishing. We should quit publishing them." Or "Zombie ideas about meta-analysis in ecology". Or

"Ecologists should all be forced to study the same question using the same methods". 😊

I don't go so far out of my way to be provocative and snarky any more. Which I think is for the best, overall. Though we know from old polls that a minority of readers would disagree. And there's a small part of me that does miss 2011 me. (And yes, I recognize that there are probably some ecologists who don't think 2011 me ever went away, and think I'm still too provocative and snarky. I respect that view even though I don't agree with it. There's no pleasing everyone as a blogger.)



Mark Vellend

on [February 10, 2021 at 2:57 pm](#) said:

Setting aside the question of how the research enterprise might be structured differently, I guess I'm thinking that it is well worth systematically assessing the state of knowledge when 50 or more studies have been done on a given topic. So if we look backward from a given point in time, if we don't do a meta-analysis, what do we do? Qualitative systematic reviews? Or, don't bother trying to pretend there is a "state of knowledge" because the evidence base is too thin for that? I just skimmed your alternate pathways to generality, and none of them (as far as I can tell) involve doing *anything* with those 50 existing studies, but rather different activities that don't involve looking at an accumulation of case studies.



Jeremy Fox

on [February 10, 2021 at 3:17 pm](#) said:

One possibly-pedantic note: if we have 50+ primary ecological *studies* (by which I mean papers), then we probably have more like 100 effect sizes. Because many primary studies report multiple effect sizes (e.g., same experiment conducted on each of several different species). If we have 50 effect sizes, that probably means we have something like 25ish published studies.

I do think that, if you wanted to argue that a quantitative summary of 50+ effect sizes is worth publishing, then your argument is the right one. We should quantitatively summarize those 50+ studies because that's often as much of an evidence base as we usually have to go on in ecology. May as well learn what we can from that evidence base. I think that's a good argument, so long as it's made in full knowledge that "learn what we can from quantitatively summarizing 50 effect sizes" might well mean "not learning much". We shouldn't, for instance, cite a meta-analysis of 50 effect sizes of X on Y as showing definitively that X typically has a huge effect on Y. As another example, when reviewing subsequent studies of the effect of

X on Y, reviewers shouldn't say things like "This study must be wrong. Meta-analysis Q shows that X has a huge mean effect on Y. So any study that fails to find an effect of X on Y must be mistaken." Etc.

But I'm also going to stick to my guns here and suggest that it would be good for the field of ecology as a whole if ecologists who saw a meta-analysis based on 50 effect sizes reflexively thought to themselves "Hmm, that's an awfully small meta-analysis. I don't know that I really buy any of its conclusions. Not with much confidence anyway."



Raphaël Proulx

on February 10, 2021 at 4:47 pm said:

Jeremy, I'd like to turn the your question around: Should we wait until we have spent our time and money on 100+ studies to conclude to a useless theory? 😊

If, based on 30 studies, someone can say that the effect of X on Y is generally positive and strong; well there are good reasons to believe that adding 30 more should not change this conclusion in a drastic fashion. On the other hand, if the conclusion is that the effect of X on Y is highly variable in direction and overall weak, there are good reasons to believe that context (moderators) is important, or that the underlying theory is plain wrong. This is usually how I read meta-analyses. I tend to put little emphasis on the exact value of the grand mean. In fact, I see meta-analyses as a great tool to debunk zombie theories and emphasise promising research avenues.



Jeremy Fox

on February 10, 2021 at 5:04 pm said:

"If, based on 30 studies, someone can say that the effect of X on Y is generally positive and strong; well there are good reasons to believe that adding 30 more should not change this conclusion in a drastic fashion. "

As long as we're confident there isn't some form of publication bias or a file drawer problem! (Which I actually think we probably should be fairly confident of in most cases in ecology. Just throwing it out there. 😊)

"In fact, I see meta-analyses as a great tool to debunk zombie theories and emphasise promising research avenues."

Testing theory is a fairly rare use of meta-analysis in ecology. Only a minority of meta-analyses are conducted as tests of theory. Which isn't to diminish the importance of those meta-analyses at all. It's just to say that there are other reasons for conducting meta-analyses that are worth thinking hard about.

Re: meta-analyses as a way to emphasize promising avenues of research: I would like someone to point out cases in which a meta-analysis paper said "we need more studies of X" or "we need more effects sizes measured in context Y", and an appreciable number of ecologists (not just one or two) went out and did that. Because I don't feel like we're short on gaps in our knowledge, or on calls for future research! I feel like, no matter what research question you want to work on ecology, you could cite *somebody* who's called for more work on that research question. Or you could identify *some* gap in the literature that provides a motivation for you asking that research question. It's not clear to me that meta-analyses *specifically* really do all that much to shape ecologists' collective research efforts. But maybe I'm wrong! Y'all can tell me: somebody name some cases where a substantial amount of research on topic or question X was conducted because a meta-analysis called for it.

I would be super-interested to hear about such cases. Because when I think of cases in which lots of people went out and researched topic X because they were inspired or influenced by some foundational paper, that paper *wasn't* a meta-analysis. It was a primary study. The biodiversity-ecosystem function bandwagon, for instance, didn't get rolling because of a meta-analysis. It got rolling mostly because of a couple of high-profile primary research studies: Tilman & Downing 1994 and Naeem et al. 1994. And when I think of the rare review or "synthesis" papers that influenced the direction of lots of subsequent ecological research, they were *not* meta-analyses. I have an old post on this:

<https://dynamicecology.wordpress.com/2020/05/25/what-are-the-most-important-influential-review-papers-in-the-history-of-ecology/>



Raphaël Proulx

on February 15, 2021 at 3:49 pm said:

As a follow up. When teaching conservation biology to undergrads, I warn them of using induction principles as evidence for taking actions. One induction principle I discussed with them last week is: N enrichment decreases plant species richness. What are the actual evidences for this statement? How bad is it? How likely is it to happen? We went on to screen the

recent literature on this topic and found the following meta-analysis:

<https://onlinelibrary.wiley.com/doi/pdf/10.1111/GEB.12856>

We concluded that:

- i) the negative effect of N enrichment is dose-dependent and both species richness and abundance are increasingly affected by increasing N inputs.
- ii) N enrichment typically makes a few winners but many more losers.
- iii) the effects were mediated by temperature and or precipitations, suggesting that mineralization and nitrification rates are key factors.
- iv) the effects were NOT mediated by ecosystem type, soil pH and fertilizer type, suggesting that the principle is general.

This meta-analysis gathered data on a total of 115 experiments (85 studies). And, yes, I am happy they published it. We concluded that the induction principle is (so far) supported, but very much dose-dependent. Unexpectedly, we also concluded that N enrichment effects are, to a great extent, independent of ecosystem type and soil conditions.

One last thing I value about meta-analyses and should be emphasized more is that the analyst has to go through a rigorous process of selecting studies (e.g. PRISMA chart). For instance, the study design must be clear. The effect size must be associated to a measure of error. It highlights methodological gaps, such as the proportion of studies that failed to provide soil pH values, for instance.

So, I think that ecologists are right when saying that a typical ecological meta-analysis should include 50ish effect sizes, but that you are right that this is probably not enough to obtain a stable grand mean effect, depending on which criterion is used to define “stable”.

All in all, I'm not sure there is a need here for bringing back the Jeremy of 2011 😊

Best -RP



Jeremy Fox

on **February 15, 2021 at 4:04 pm** said:

Interesting example Raphael. Though I wonder a bit whether the example you cite is a typical one. I feel like the effect of N enrichment on terrestrial plant species richness is one of the best studied, most general, most repeatable effects in community ecology.

Re: “a rigorous process of selecting studies”, I feel like that’s pretty standard in ecological meta-analyses these days. Though there’s a part of me that wonders if you can find any statistical signature that distinguishes (say) meta-analyses that include a PRISMA chart from those that don’t. I’m being a bit cheeky here, obviously. I think it’s good for meta-analyses to have clear study selection criteria, even if one can’t detect any systemic differences in results between meta-analyses that have clear study selection criteria and meta-analyses that don’t. Still, I do wonder a little.



Brian Cade

on **February 10, 2021 at 5:10 pm** said:

Perhaps the mistake being made in all these meta analyses is thinking that a mean effect size has any useful interpretation given the inherent heterogeneity in these studies. An interval of effect sizes, say 5th to 95th percentiles, makes more sense to me and more effectively captures the variation actually being estimated. There is no single “effect size”, there are only intervals of “effect sizes”. Unfortunately, most of the statistical procedures being used such as mixed-effects models propels people towards computing mean effect sizes.



Jeremy Fox

on **February 10, 2021 at 5:14 pm** said:

Sure. Although if the mean isn’t meaningful, would either a 95% c.i. around the mean, or a 95% prediction interval for the effect sizes (“95% of effect sizes will fall in this range”), be any more meaningful? Like, if you shouldn’t be averaging together all these effect sizes, should you be calculating their variance either? Or even plotting them all on the same graph?

I’ll take this opportunity to shamelessly re-up this old post:

<https://dynamicecology.wordpress.com/2015/05/14/in-a-variable-world-are-averages-just-epiphenomena/>



Jeff Houlahan

on **February 11, 2021 at 10:28 am** said:

I'm with Jeremy on this – further, this is an argument that means aren't useful if there is inherent heterogeneity. The only time we use means is when there is inherent heterogeneity. If there is no heterogeneity there is no need for a mean. Presumably, the argument is that – above some level of heterogeneity, means have no value. It's not obvious to me that's true. But if it is – how do we decide on the 'critical' level of heterogeneity?



Jeremy Fox

on February 11, 2021 at 10:41 am said:

One principled way to decide on whether there's "too much" heterogeneity is to ask whether additional observations any additional "indirect" information about your sample statistics. Efron (1977) is great on the basic idea behind this: <http://statweb.stanford.edu/~ckirby/brad/other/Article1977.pdf>



Brian McGill

on February 11, 2021 at 1:46 pm said:

We Brian's have to stick together! I'm with Brian on this. The point is there is a lot of heterogeneity across systems. So trying to find a mean by boiling that down to a single number is misleading (there may be no system that actually has that mean effect size). But if you have any comparative goals (i.e. goals beyond a single system), then quantifying the percentile ranges (not confidence intervals) or more completely an empirical histogram or probability distribution is the cross-system summarization of that heterogeneity. Just like a body size distribution is an accurate, non-simplistic representation of the "body size" of a community. Giving a mean body size is useless and misleading. Giving a distribution of body sizes is accurate but also informative and predictive of the body size of the next individual sampled. This is also explanatory of the main point of this blog post – accurately estimating a probability distribution requires much more data than estimating a mean.



Jeremy Fox

on February 11, 2021 at 1:59 pm said:

"This is also explanatory of the main point of this blog post – accurately estimating a probability distribution requires much more data than estimating a mean."

I like that way of putting it.

That way of putting it sidesteps some of the conceptual issues about heterogeneity—whether some numbers shouldn't be lumped together onto the same graph, or into the same statistical analysis, because they're "too heterogeneous". Whatever numbers you've decided comprise your sample, you can think of them as having been sampled from *some* probability distribution or other. If you say you don't really care about the mean of that distribution, you care about its variance or its overall shape, fair enough. But then you have to bite the bullet and recognize that you're going to need a frickin' massive sample size to estimate the features of the distribution that you care about.

This is useful because it's inspired me to go back to my big compilation of all the effect sizes from 476 ecological meta-analyses and look at the prediction intervals for the data (as opposed to the confidence interval for the mean). How wide is, say, the central 50% of the observed distribution of effect sizes? The central 95%? What fraction of effect sizes are positive, and what fraction are negative? And how do all those things change over time as more effect sizes are published? How many effect sizes do you have to have before you can be reasonably confident that (say) you've pinned down the location of the central 50% of the distribution of effect sizes?



Jeremy Fox

on **February 10, 2021 at 6:51 pm** said:

Somewhat off topic: how come meta-analysis has been taken up so enthusiastically in ecology, but not so much in other fields? For instance:

Adam J Calhoun · Feb 10, 2021



@neuroecology · [Follow](#)

Are there any metanalyses in systems neuroscience?

Adam J Calhoun

@neuroecology · [Follow](#)

broader question: why so few? what does systems neuro have against meta analysis?

Do we just agree on everything? Or is everyone just doing totally different experiments

9:17 PM · Feb 10, 2021



4



[Reply](#)



[Share](#)

[Read 5 replies](#)



Jeremy Fox

on **February 10, 2021 at 6:53 pm** said:

And if the answer to Adam Calhoun's question ("Why so few meta-analyses in neuroscience?") is given by the tweet below, then maybe neuroscientists and ecologists should talk. 😊

Adam J Calhoun · Feb 10, 2021



@neuroecology · [Follow](#)

Replying to @neuroecology

broader question: why so few? what does systems neuro have against meta analysis?

Do we just agree on everything? Or is everyone just doing totally different experiments

Jim Bohnslav

@jbohnslav · [Follow](#)

too many variations in analysis, and slightly different experimental protocols that could conceivably have large effects

9:22 PM · Feb 10, 2021



8



[Reply](#)



[Share](#)

[Read 2 replies](#)



Jeremy Fox

on [February 10, 2021 at 6:56 pm](#) said:

Thinking about it for another 10 seconds, I've decided my joking remark in the previous tweet isn't a joke. It would actually be super-interesting to have a conversation between folks in field X, who have taken up approach Y in a big way, and folks in some other field who haven't taken up approach Y *for reasons that would seem to apply to field X as well*.



Neva Knott

on [February 10, 2021 at 8:25 pm](#) said:

Great post.



Jeremy Fox

on [February 10, 2021 at 8:49 pm](#) said:

Thanks!

**Jeremy Fox**on **February 11, 2021 at 12:33 am** said:

Had forgotten this old post, which I'll re-up as it's relevant:

<https://dynamicecology.wordpress.com/2017/03/29/how-far-can-the-logic-of-shrinkage-estimators-be-pushed-or-when-should-you-compare-apples-and-oranges/>

**Oscar Miles**on **February 18, 2021 at 2:33 pm** said:

I attended a seminar course about ecological meta-analyses in 2011. We were clobbered over the head with the issue of effect size- spending about 1/3 of the course on this very issue. I do not know how common it is to offer coursework for ecological meta-analyses, or how curricula might vary among programs, but it could be a lack of education accounts for what you've observed. I agree with your tact that meta-analyses are of limited value in ecology. I think they are most suitable for designing new studies, as they can provide a "course grained" indication of what might be happening in nature.

Pingback: [Best examples of moderator variables explaining variation in effect size in ecological meta-analyses? | Dynamic Ecology](#)

This site uses Akismet to reduce spam. [Learn how your comment data is processed](#).