



Symposium Article

The Heritability of Behavior: A Meta-analysis

Ned A. Dochtermann, Tori Schwab, Monica Anderson Berdal*,
Jeremy Dalos*, and Raphaël Royauté*

From the Department of Biological Sciences, North Dakota State University, 329 Stevens Hall, Fargo, ND 58104 (Dochtermann, Schwab, Anderson Berdal, Dalos, and Royauté).

Address correspondence to N. A. Dochtermann at the address above, or e-mail: ned.dochtermann@gmail.com.

Received August 31, 2018; First decision February 5, 2019; Accepted March 27, 2019.

Corresponding Editor: Anne Bronikowski

*These authors contributed equally and are listed in alphabetical order.

Abstract

The contribution of genetic variation to phenotypes is a central factor in whether and how populations respond to selection. The most common approach to estimating these influences is via the calculation of heritabilities, which summarize the contribution of genetic variation to phenotypic variation. Heritabilities also indicate the relative effect of genetic variation on phenotypes versus that of environmental sources of variation. For labile traits like behavioral responses, life history traits, and physiological responses, estimation of heritabilities is important as these traits are strongly influenced by the environment. Thus, knowing whether or not genetic variation is present within populations is necessary to understand whether or not these populations can evolve in response to selection. Here we report the results of a meta-analysis summarizing what we currently know about the heritability of behavior. Using phylogenetically controlled methods we assessed the average heritability of behavior (0.235)—which is similar to that reported in previous analyses of physiological and life history traits—and examined differences among taxa, behavioral classifications, and other biologically relevant factors. We found that there was considerable variation among behaviors as to how heritable they were, with migratory behaviors being the most heritable. Interestingly, we found no effect of phylogeny on estimates of heritability. These results suggest, first, that behavior may not be particularly unique in the degree to which it is influenced by factors other than genetics and, second, that those factors influencing whether a behavioral trait will have low or high heritability require further consideration.

Subject: Quantitative genetics and Mendelian inheritance

Keywords: behavior, heritability, phylogeny

Understanding the ability of a phenotypic trait to evolve in response to selection requires, among other things, knowledge of the degree to which there is genetic variation present in that trait. For labile traits, that is, traits which are expressed repeatedly and variably in response to internal or external stimuli, estimating the *relative* magnitude of genetic variation present is particularly important as these traits are often strongly affected by the environment. Questions

about the relative magnitude of genetic variation can be addressed by estimating the heritability—the magnitude of additive genetic variation present in a trait relative to the total phenotypic variation (i.e., b^2)—of the traits of interest (Lynch and Walsh 1998). A high heritability indicates that much of the phenotypic variation is attributable to genetic differences amongst individuals. Importantly, a high heritability is not necessarily indicative of high evolutionary

potential due both to a lack of independence between additive genetic and environmental variation and due to an expected negative correlation between heritability and selection differentials (Houle 1992; Hansen et al. 2011).

Behavior, in particular, is frequently assumed to be highly plastic (West-Eberhard 2003), suggesting that the magnitude of genetic variation relative to the effects of the environment on behavior should be low. Unfortunately, the validity of this assumption is unclear but would be supported if the average heritability of behavior is low relative to other labile traits like physiology and life history. Understanding general patterns in the heritability of behavior is also important as the validity of evolutionary inferences drawn in behavioral ecology depends on there being underlying heritability in traits of interest, that is, “the phenotypic gambit” (Grafen 1984). Specifically, behavioral ecologists often assume that differences among groups (e.g., age/stage classes, sexes, populations, and species) are adaptive and underpinned by genetic differences.

Relevant to these assumptions there have been several previous efforts reviewing the heritability of behavior (Table 1). Besides 2 studies which indirectly summarized a sample of behavioral heritabilities (Dochtermann 2011; Dochtermann et al. 2015), 3 key sources summarize what we currently know about the heritability of behavior: 1) In a review of behavioral heritabilities, Mousseau and Roff (1987) found that behaviors had heritabilities similar to those reported for physiological traits. Likewise, 2) Stirling et al. (2002), in the most expansive review thus far, compared heritabilities among behavioral, life history, and morphological traits and found that there was not a significant difference in heritability between behavior and life history. More recently, 3) van Oers and Sinn (2013) reviewed heritabilities of “personality traits” and found that these behaviors had a raw average of 0.26 (Table 1).

Taken together these results suggest that behaviors are comparable to other non-morphological traits in the relative contribution of additive genetic variation to phenotypic variation. As discussed above, behavior exhibiting heritability comparable to that of physiology and life history might be surprising because animal behaviorists and behavioral ecologists often assume that behavior is more labile than other types of traits. While behavior is indeed often labile, these results suggest that it is not especially so when compared to physiology and life history. A general finding of all 3 studies was also that heritability was higher when measured in the field versus in the lab. This could be due to several non-mutually exclusive explanations including 1) the possibility that permanent environmental effects are conflated with genetic variation in the field and 2) rearing in laboratory conditions reduces genetic variation (e.g., via founder effects and subsequent bottlenecks in laboratory population sizes).

Besides these 3 reviews of behavioral heritabilities, another potentially useful source of information about the heritability of behavior

might be provided by estimates of repeatability. Repeatability and heritability are mathematically and conceptually related (Boake 1989). In particular, the numerator of repeatability—that is, among-individual variation—combines genetic and permanent environment sources of variations, while only additive genetic variation is considered in the calculation of heritability (Dingemanse and Dochtermann 2014; Dochtermann et al. 2015). As such, estimates of repeatability will typically set an upper limit to heritability and therefore provide estimates of maximum likely heritabilities (but see Dohm 2002 for important theoretical exceptions). The best source of information regarding behavioral repeatabilities is provided by an extensive meta-analysis carried out by Bell et al. (2009).

In their meta-analysis Bell et al. (2009) examined whether 1) some types of behaviors tended to be more repeatable than others, 2) which taxa exhibited higher repeatabilities, 3) if time intervals between observations influenced repeatability, 4) if the number of measurements per individual influenced repeatability, and 5) if factors such as sex, age, and environment influenced repeatability estimates. These authors determined that aggression, mating behavior, and habitat selection were the most repeatable. Bell et al. (2009) also found that endotherms and males had a tendency to show higher repeatabilities. Finally, behaviors measured in the field tended to exhibit higher repeatabilities, as also observed for heritabilities (Mousseau and Roff 1987; Stirling et al. 2002; van Oers and Sinn 2013).

Despite the 3 prior reviews of heritability estimates, concerns about the robustness of their estimates remain. First, because neither Mousseau and Roff (1987) nor Stirling et al. (2002) conducted formal meta-analyses, their general conclusions conflate estimation error with other sources of variation. Second, none of the 3 prior reviews incorporated phylogeny in their analyses and so the possibility of taxonomic differences in average heritability was not assessed. Here we sought to address these concerns and to specifically determine: 1) whether there are taxonomic differences in heritability, 2) if heritability differs between field and laboratory studies, 3) if heritability differs between domestic and wild animals, 4) whether heritability estimates differ between invertebrates and vertebrates, 5) whether ectotherms and endotherms exhibit differences in heritability, and 6) whether heritabilities differ among behavioral categories.

Methods

Dataset Assembly

Heritability estimates were gathered via Web of Science searches of the following journals: American Naturalist, Animal Behaviour, Behavioral Ecology, Behavioral Ecology & Sociobiology, Behavior Genetics, Behaviour, Ecology, Ethology, Evolution, Heredity,

Table 1. Previous estimates of behavioral heritabilities

Authors	Year	Mean heritability	Number of estimates	Inclusion/exclusion criteria
Mousseau & Roff	1987	0.30	105	Wild, outbred stock
Stirling et al.	2002	0.31	454	All estimates except from diallelic crosses
Dochtermann	2011	0.31	200	Genetic correlations reported
van Oers & Sinn	2013	0.26	209	“Personality traits”
Dochtermann et al.	2015	0.14	70	Both heritability and repeatability were reported
Weighted average		0.29	1038 ^a	

^aMany of the datasets for these papers included estimates also used in prior analyses, thus this sample size and the associated weighted average is pseudo-replicated.

Proceedings of the Royal Society, B. For behavioral journals we used the key word “heritab*”, while for evolutionary ecology journals we used the key words “heritab* AND behav*”. We also restricted our search to nonhuman studies by including “NOT twin NOT human” in all searches. We searched these journals using these terms to locate all indexed articles published between January 2000 to May 24, 2018. January 2000 was used as a starting date as [Stirling et al. \(2002\)](#) reviewed estimates through December 1999 and our search thus represents an independent review of the literature since that time.

Studies that gave estimate ranges (e.g., >.5), had heritabilities greater than 1 or less than zero, or that were based on human studies were excluded from the dataset (as in [Dochtermann 2011](#); [Dochtermann et al. 2015](#)). Also excluded from the data set were estimates of morphological, physiological, and life history traits. Published studies identified according to these criteria were then analyzed to determine whether species were “domestic,” “semi-domestic,” or “wild,” whether heritability was estimated in laboratory or field conditions, taxonomic classification, and the behaviors measured classified to one of 11 categories ([Table 2](#)). These categories represent a reduced set of those used by [Bell et al. \(2009\)](#). For example, here we grouped all mating behaviors together in this study whereas [Bell et al. \(2009\)](#) more finely distinguished among categories. Species used for economic gain were classified as domestic, species maintained in captivity for a large number of generations were classified as semi-domestic, and all others were classified as wild ([Stirling et al. 2002](#)). Following [Stirling et al. \(2002\)](#), studies done in a species’ “natural habitat” were classified as field studies and studies conducted under controlled conditions were classified as laboratory studies.

Dataset Characteristics

Our search initially identified 639 publications. According to the above criteria, this sample reduced down to 101 publications and

Table 2. Behavioral categories used in the meta-analysis and corresponding operational definitions

Behavior	Definition
Activity	Movement rate and patterns by individuals
Aggression	Antagonistic behavior directed towards another individual.
Boldness	Behavior involving predator avoidance and inspection. Includes measures of “antipredator” response such as escape behavior or latency to emerge from protected areas.
Communication	The transfer of information from one individual or group to another.
Exploratory behavior	An individual investigating a novel environment.
Foraging	Behaviors involved in searching for, handling, and consuming food.
Mating	Mate search, courtship, and copulation behaviors.
Migration/dispersal	Movement among areas either repeatedly or on one occasion during an organism’s life
Parental effort	Parental expenditure that benefits the offspring.
Social behavior	Affiliative and grouping behaviors (e.g., allogrooming or shoaling).
Other	Behaviors not included in the above.

a total of 476 estimates of heritability ([Figure 1](#)). Heritabilities were most frequently reported for mating behavior, boldness, and activity. There were also a large number of estimates classified as “other behaviors” (167). However, 157 of these were from 2 studies of nonhuman primates and were a variety of psychological assays. Taxonomically the dataset was heavily biased, with 84 percent of estimates from Chordates.

Data Analysis

To assess our 6 a priori questions we conducted a formal meta-analysis using the metafor package for mixed effect, phylogenetically controlled, models ([Viechtbauer 2010](#)). These analyses assume that residual variation is normally distributed. Because heritability estimates are, by definition, not normally distributed, we used Fisher’s z -transformation (Zr) prior to all analyses and back-transformed parameter estimates to heritabilities (h^2) for interpretation. h^2 estimates are reported with 95% confidence intervals for moderators (e.g., Behavioral Category). While Zr was still not normally distributed due to h^2 never being less than zero, this transformation best conformed to meta-analytic requirements (see also: [Bell et al. 2009](#), [van Oers and Sinn 2013](#)).

We first constructed a phylogenetic meta-regression model with only an intercept but no other fixed effects (i.e., moderators). Article identity and the species to which the study organism belonged were included as random effects. Evolutionary nonindependence was modeled using a phylogenetic tree constructed using NCBI molecular data via phyloT ([Figure 2](#), [Letunic 2015](#)). From this model we calculated 1) sampling variance following [Nakagawa and Santos \(2012\)](#), 2) the variation attributable to the published article an estimate was taken from, 3) variation attributable to phylogeny, and 4) calculated the total heterogeneity in the dataset (I^2). Estimation of the variation attributable to phylogeny addressed the first of our questions, whether there were taxonomic differences in heritability.

After this global model we tested the remainder of our questions via the inclusion of biological moderators. Specifically, we evaluated a statistical model incorporating the same random effects structure as above but also including the categories of field versus laboratory, wild versus semi- and domestic, vertebrate versus invertebrate, and endotherm versus ectotherm (all members of the classes Aves and

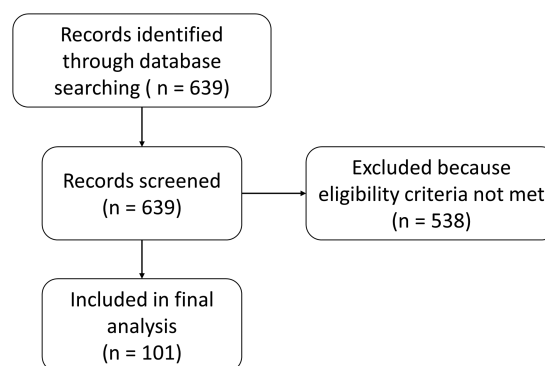


Figure 1. PRISMA diagram describing the included studies. Of the 639 studies identified via database searches, only 101 were ultimately included in the final analysis. Most of the 538 excluded studies did not report heritabilities, despite heritability being discussed. As our eligibility criteria included particular data requirements (i.e., no reports of heritability ranges), this PRISMA diagram does not include a separate branch for studies excluded due to insufficient data. The 101 studies included in the analysis yielded 476 total estimates of heritability.

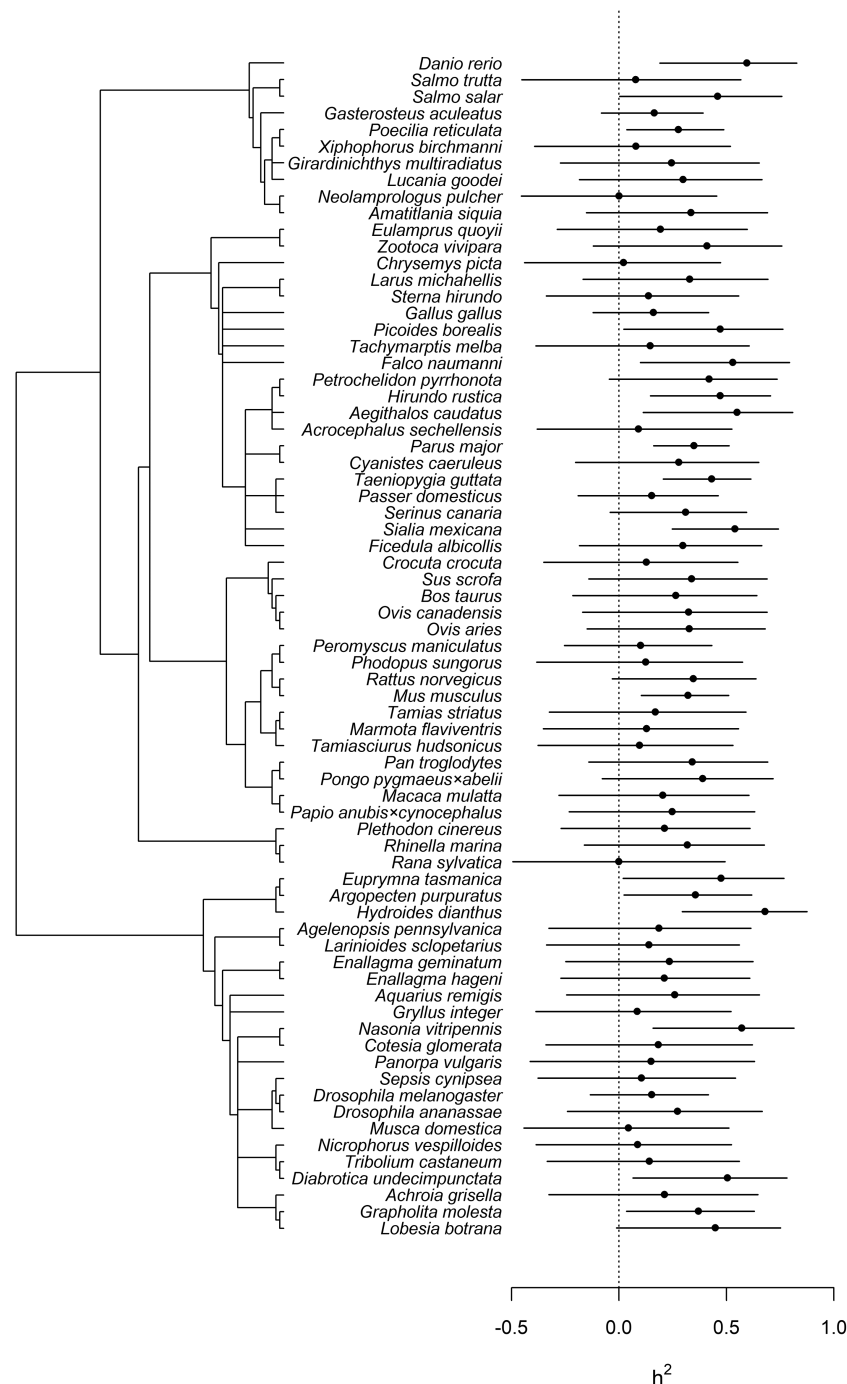


Figure 2. Phylogenetic tree (left) and associated forest plot (right) for mean effect sizes ($\pm 95\%$ confidence intervals) by species. h^2 was back-calculated from Z_r values and thus can have confidence intervals that overlap zero.

Mammalia were categorized as endotherms). While endotherm and ectotherm are broad categories, we used them here due to the expectation that animals categorized as ectotherms might be expected to exhibit greater influences of the environment on behavioral expression, resulting in lower heritabilities. To this model, we also added behavioral category (Table 3). While meta-analyses often estimate the effects of moderators such as these in a pairwise manner (i.e., an intercept-only model vs. a model only including the moderator of interest), here we tested the influence of each moderator versus a full model as otherwise conflation of moderator categories can lead

to inappropriate estimates of effect sizes. We assessed the significance of each fixed effect based on likelihood ratio tests of the moderator of interest.

To address potential misestimation due to Z_r not actually being normally distributed we estimated the global average heritability ($\pm 95\%$ credibility intervals) using mixed-effects Beta regressions using the original h^2 values and fit using Bayesian estimation via the brms package (Bürkner 2017). Beta regressions were not used for the full set of analyses because they do not allow the inclusion of fixed variances, as is necessary for meta-analyses.

As an additional caveat to our general analyses, reported sample sizes across articles were often unclear because of nonindependence of individuals (Noble et al. 2017). Following Noble et al. (2017) we used the conservative approach of using the sample size at the highest grouping levels (e.g., the number of families rather than individuals measured). Unfortunately, due to inconsistent reporting, this was not always possible or clear. Therefore, we also provide redundant analyses in the [Supplementary Material](#) wherein all analyses are conducted as reported in the main text but with sample sizes set to the smaller of 2 values (100 or the reported number). This analysis led to the same general conclusions vis-à-vis phylogenetic signal and the effect of behavioral category ([Supplementary Tables S1–3](#)).

Finally, we tested for publication bias using Egger's regression tests (Egger et al. 1997), inspection of funnel plots, and trim-and-fill methods to estimate the number of studies that may have been missing from our dataset. All analyses and data are provided as [Supplementary Material](#).

Results

Overall Average Heritability of Behavior

The average heritability of behavior was 0.235 (95% credibility interval: 0.200: 0.271) as estimated via Bayesian Beta regression. Dataset heterogeneity as estimated via restricted estimate maximum likelihood was high ($I^2 = 99.78$, [Table 3](#)), justifying the examination of biological influences on the magnitude of heritability. Further investigation of sources of heterogeneity also revealed that phylogeny did not contribute to the observed heterogeneity in estimates of heritability ([Table 3](#)), a finding supported by a lack of apparent patterns across the species sampled ([Figure 2](#)).

Effects of Biological Moderators on Estimates of Heritability

Four of the 5 included biological moderators did not systematically influence heritability. Heritability did not significantly differ between

Table 3. Magnitudes of variation (i.e., heterogeneity) due to article ID, phylogenetic structure, sampling error, and that unattributable to any modeled source (residual)

Source of variation	%
Article	41.51
Phylogeny	0
Residual	57.95
Sampling error	0.22
I^2	99.47

I^2 is the proportion of total variation not attributable to sampling error.

invertebrates and vertebrates ([Table 4](#), [Figure 3](#), $\chi^2 = 1.92$, $df = 1$, $P = 0.17$) nor among semi-domestic, domestic, or wild species ([Table 4](#), [Figure 3](#), $\chi^2 = 0.47$, $df = 2$, $P = 0.26$). Similarly, whether a species was classified as an ectotherm or endotherm did not significantly affect heritability ([Table 4](#), [Figure 3](#), $\chi^2 = 3.57$, $df = 1$, $P = 0.06$). Lastly, whether estimation was conducted in field or laboratory conditions also did not significantly influence heritability ([Table 4](#), [Figure 3](#), $\chi^2 = 0.47$, $df = 1$, $P = 0.49$).

In contrast, considerable and significant variation was also observed among behavioral categories ([Table 4](#), [Figure 3](#), $\chi^2 = 189.41$, $df = 10$, $P < 0.001$). The highest average heritability was observed for migratory and dispersal behaviors ([Figure 3](#), $h^2 = 0.456$ [95% confidence interval: 0.331: 0.565]), although only 15 estimates informed this average. Communicative and mating behaviors also exhibited relatively high heritabilities ($h^2 = 0.351$ [95% confidence interval: 0.229: 0.462] and 0.347 [95% confidence interval: 0.269: 0.421], respectively) though communicative behaviors were poorly represented in the dataset ([Figure 3](#)). Most other behaviors exhibited average heritabilities ranging from around 0.26 to 0.29 with the exception of foraging behaviors ($h^2 = 0.196$ [95% confidence interval: 0.136: 0.254]) and “other behaviors” ($h^2 = 0.146$ [95% confidence interval: 0.091: 0.200]) which exhibited particularly low average heritabilities ([Figure 3](#)).

Publication Bias

Egger's regression results suggest that the dataset lacks publication bias ($t_{474} = -1.42$, $P = 0.16$). However, this test was necessarily conducted on residuals of a model that ignored the phylogenetic structure and the nestedness of estimates within studies and so may be anti-conservative. Unfortunately, funnel plot inspection ([Figure 4](#)) is ambiguous to this topic, given that Z_r necessarily cannot extend below zero for our dataset, but does suggest some publication bias with the largest heritabilities estimated in studies with lower precisions. Nonetheless, and consistent with Egger's regression results, a trim-and-fill analysis estimated the number of missing low heritabilities as zero (SE: 9.31).

Discussion

Our estimate of the global average of heritability for behavior, 0.235, is slightly lower than previous reports, for which a weighted average of 0.29 was estimated ([Table 1](#)). This discrepancy is possibly due to several non-mutually exclusive factors including that previous reports estimated heritability under the assumption that it was normally distributed rather than Beta distributed and large uncertainties around previous reports. Given the large uncertainty around estimates of heritability in general, we do not consider the difference in estimates reported here to be biologically substantive.

Table 4. Significance test results for the inclusion of particular moderators

Moderator	Log-Likelihood	χ^2	df	Test dfs	P-value
Endo- vs. ectotherm	-2229.10	3.57	17	1	0.06
Invertebrate vs. vertebrate	-2228.28	1.92	17	1	0.17
Behavioral category	-2322.05	189.41	8	10	<0.001
Semi- vs. domestic vs. wild	-2228.67	2.72	16	2	0.26
Field vs. laboratory	-2227.55	0.47	17	1	0.49
Full model	-2227.31		18		

Only the inclusion of behavioral categories significantly improved model fit, indicated in bold.

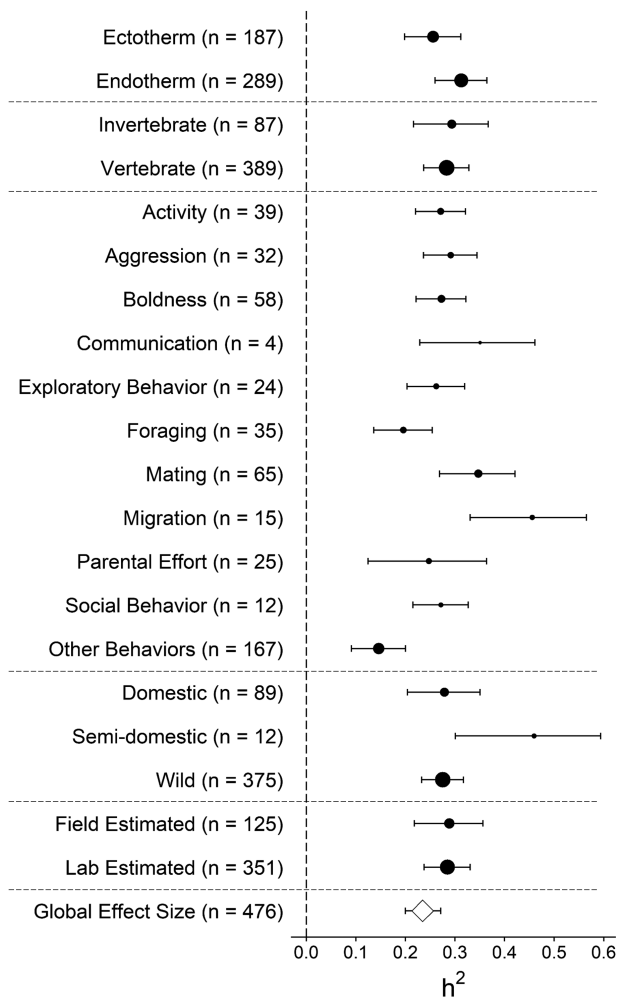


Figure 3. Forest plot of estimated mean effect sizes ($\pm 95\%$ confidence intervals) for each moderator. Point size is proportional to the sample size for a particular moderator.

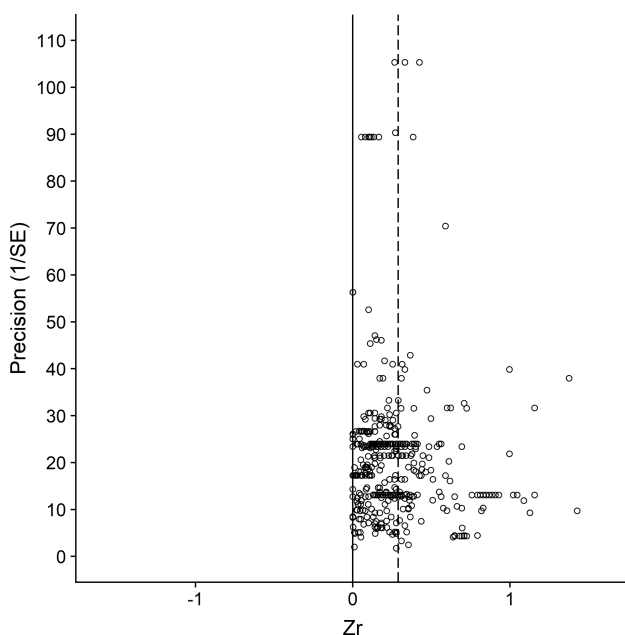


Figure 4. Funnel plot for the estimation of publication bias in the full data set.

Our results are also informative in several novel ways: First, by specifically incorporating phylogenetic structure in our analysis we were able to both control for evolutionary nonindependence and estimate phylogenetic contribution. Second, rather than acting as an update of previous reviews, because we did not include the studies included in Mousseau and Roff (1987) nor in Stirling et al. (2002), our results act as an independent assessment of prior conclusions. Third, the inclusion of particular biological moderators allows broad inferences to be drawn within a formal meta-analytic framework.

Our first *a priori* question was whether there were taxonomic patterns present in the data. Based on the percent of heterogeneity present in the data that was attributable to phylogeny (0%, Table 3), this is strongly suggested to not be the case. Examination of species-level estimates (Figure 1) and the lack of a significant difference between the broad taxonomic categories of vertebrates and invertebrates (Figure 2, Table 4) provides additional support for the inference that heritabilities do not vary systematically among taxa. Given existing arguments in the literature that reduced heritabilities are indicative of a particular trait being under greater selection (Mousseau and Roff 1987 but see Houle 1992), this result might suggest that behaviors have not been under stronger selective pressures in some taxa than others. However, a more thorough investigation of this question would require the estimation of trait evolvabilities (Houle 1992).

As with phylogenetic signal, we found no support for heritabilities differing between laboratory and field conditions, nor among wild, domestic, or semi-domestic species. These results are particularly interesting as Stirling et al. (2002) found that heritability of behavior was highest for domestic species and lowest for wild species, a pattern we had expected would be recreated. Our finding that heritability did not differ if estimated in the laboratory or under field conditions is also interesting as Bell et al. (2009), found that field estimated repeatabilities were higher than those estimated in the laboratory. Given the mathematical and biological relationship between heritability and repeatability (Boake 1994, Dochtermann et al. 2015), if the studies included in our meta-analysis and the meta-analysis conducted by Bell et al. (2009) provide unbiased random samples of relevant estimates, these contradictory findings could arise from one of 2 mechanisms: 1) $G \times E$ effects wherein realized genetic variance is greater under field conditions or 2) permanent environmental effects have a greater contribution to behavior under field rather than laboratory conditions. These mechanisms are not mutually exclusive and both seem equally likely.

Behavioral category had the strongest effect on variation in heritability estimates. In our analysis, migratory and mating behaviors exhibited particularly high average heritabilities, while foraging and “other behaviors” exhibited low average heritabilities (Figure 3). Some specific category averages were much different than those previously reported. For example, Stirling et al. (2002) reported an average heritability of 0.29 for foraging behaviors whereas here the estimated average was 0.18. Our estimate of a lower heritability makes intuitive sense as foraging behaviors are likely to be highly state dependent and thus strongly influenced by environmental variation. Consistent with this rationale, Stirling et al. (2002) also reported a high coefficient of residual variation of foraging behaviors. Also differing from our findings, van Oers and Sinn (2013) reported an average heritability for aggression of 0.14 whereas our estimate here was much higher: 0.29. While it is not clear why these 2 estimates differ, it is important to note that van Oers and Sinn (2013) used very different search terms than those employed here. Nonetheless, our search terms should have identified many of the same studies. More generally, the broad variability in heritability

averages among behavioral categories requires further theoretical consideration as to what types of behavior might be expected to be more influenced by genetic or environmental contributors.

Two main caveats should be considered in interpreting these data: First, we did not calculate “evolvability” for our dataset (Houle 1992; Hansen et al. 2011) because many studies lacked required values or were on scales where doing so would be inappropriate. Thus, our results cannot properly answer questions as to how rapidly behaviors might be expected to respond to unit selection. Second, and perhaps more important, our dataset is strongly taxonomically biased: 83% of our estimates were from vertebrates (Figure 2) despite vertebrates comprising only ~7% of described animal species (Rosenthal et al. 2017). Only 4 phyla were included in our analysis (with only one estimate for Annelida) and, even within the most well-represented phyla—Chordata—taxonomic over-representation of particular taxonomic groups was also extreme: of the 400 heritability estimates for chordates 185 were for mammals, 106 were for birds, and 91 were of ray-finned fishes. Despite the lack of phylogenetic signal in the data, this level of unequal sampling is problematic for several reasons, problems that also affected prior meta-analyses and reviews of heritabilities. First, phylogenetic signal can be difficult to detect if there is poor coverage. Second, this finding is indicative of a broader problem in the behavioral literature. As discussed by Rosenthal et al. (2017), this misrepresentation is pervasive in the study of animal behavior. The most major implication of this misrepresentation is that we may grossly misunderstand the heritability of behaviors or, worse yet, misunderstand the behavior of animals in general due to a narrow taxonomic focus.

It is also worth putting our results into the broader context of what we know about the heritability of phenotypic traits in general. The average heritability of 0.235 estimated here is not substantively different than either the estimated average heritability of life history traits (0.262) found by Mousseau and Roff (1987), particularly given the large uncertainties with which any single study can estimate heritability. Although we did not statistically test for a difference here, Stirling et al.’s (2002) heritability estimate for behavior (0.31) did not differ from estimates for life history traits (0.262) and Mousseau and Roff (1987) found no difference between their estimates of the heritabilities of physiology (0.33) and behavior (0.30). Further, while life history, physiology, and behavioral traits all exhibit lower heritabilities than morphology (0.461, Mousseau and Roff 1987), the fact that they differ little from each other should not necessarily be surprising. While behavior is often considered to be highly plastic, behavior, physiology, and some life history traits are all typically expressed multiple times throughout an organism’s life. Behaviors and physiological responses, in particular, are also often expressed in response to changing environmental conditions over very short time frames and thus should be expected to be similarly labile.

Our results also demonstrate that populations harbor substantial genetic variation in behaviors relative to that variation attributable to the environment. This is consistent with general arguments that genetic variation is abundant and thus not likely to constrain evolutionary processes (e.g., Blows and Hoffmann 2005). While considerable attention over the last 15 years in the study of behavior has been directed to understanding contributors to behavioral variation (e.g., Bell et al. 2009; Sih et al. 2015), what factors generally maintain variation remains unclear. While it has been suggested that behavior may integrate with life history and physiological traits in predictable ways with a concomitant maintenance of behavioral variation (Reale et al. 2010), recent meta-analyses suggest that these

predictions may not be generally supported (Royauté et al. 2018). However, despite a potential lack of predictable trait associations, it may still generally be the case that variation in combinations of traits is limited (Blows 2007; Hansen and Houle 2008; Dochtermann and Dingemanse 2013). Absolute and quantitative constraints on evolution imposed by genetic correlations may be one mechanism that genetic variation, like that estimated here for behavior, is maintained in populations in the face of selection and drift (Walsh and Blows 2009). Understanding why variation is maintained in behaviors may, therefore, require a focus both on multiple behaviors and multiple trait types.

Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

Funding

National Science Foundation (IOS 1557951 to N.A.D.).

Acknowledgments

We thank A. Bronikowski for helpful discussions and for the invitation to participate in the AGA President’s Symposium. We also thank S. Nakagawa and an anonymous reviewer, both of whom provided comments and recommendations that greatly improved this paper.

Data Availability

Supplementary materials, code for analyses, and data are available on GitHub (github.com/DochtermannLab/BehavioralHeritability) and archived online at Dryad (doi:10.5061/dryad.b38k42m).

Author Contributions

T.S. and N.A.D. jointly designed the project. N.A.D. conducted analyses. All authors collected data and contributed to the writing of the manuscript.

References

- Bell AM, Hankison SJ, Laskowski KL. 2009. The repeatability of behaviour: a meta-analysis. *Anim Behav.* 77:771–783.
- Blows MW. 2007. A tale of two matrices: multivariate approaches in evolutionary biology. *J Evol Biol.* 20:1–8.
- Blows MW, Hoffmann AA. 2005. A reassessment of genetic limits to evolutionary change. *Ecology.* 86:1371–1384.
- Boake CR. 1994. *Evaluation of applications of the theory and methods of quantitative genetics to behavioral evolution. Quantitative genetic studies of behavioral evolution.* Chicago: University of Chicago Press. p. 305–325.
- Boake CRB. 1989. Repeatability - its role in evolutionary studies of mating-behavior. *Evol Ecol.* 3:173–182.
- Bürkner PC. 2017. brms: An R package for Bayesian multilevel models using Stan. *J Stat Softw.* 80:1–28.
- Dingemanse NJ, Dochtermann NA. 2014. Individual behaviour: behavioural ecology meets quantitative genetics. In: Charmanier A, Garant D, Kruuk LEB, editors. *Quantitative genetics in the wild.* Oxford University Press, p. 54–67.
- Dochtermann NA. 2011. Testing Cheverud’s conjecture for behavioral correlations and behavioral syndromes. *Evolution.* 65:1814–1820.
- Dochtermann NA, Dingemanse NJ. 2013. Behavioral syndromes as evolutionary constraints. *Behav Ecol.* 24:806–811.
- Dochtermann NA, Schwab T, Sih A. 2015. The contribution of additive genetic variation to personality variation: heritability of personality. *Proc R Soc B Biol Sci.* 282:20142201.

- Dohm MR. 2002. Repeatability estimates do not always set an upper limit to heritability. *Funct Ecol.* 16:273–280.
- Egger M, Smith GD, Schneider M and Minder C. 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 315: p. 629–634.
- Grafen A. 1984. Natural selection, kin selection and group selection. In: Krebs JR, Davies NB, editors. *Behavioural ecology: An evolutionary approach*. Sunderland (MA): Sinauer. p. 62–84.
- Hansen TF, Houle D. 2008. Measuring and comparing evolvability and constraint in multivariate characters. *J Evol Biol.* 21:1201–1219.
- Hansen TF, Pélabon C, Houle D. 2011. Heritability is not evolvability. *Evol Biol.* 38:258.
- Houle D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics.* 130:195–204.
- Letunic I. 2015. phyloT: Phylogenetic Tree Generator. Available from: <http://phlot.biobyte.de>.
- Lynch M, Walsh B. 1998. *Genetics and analysis of quantitative traits*. Sunderland (MA): Sinauer Associates.
- Mousseau TA, Roff DA. 1987. Natural selection and the heritability of fitness components. *Heredity (Edinb.)*. 59 (Pt 2):181–197.
- Nakagawa S, Santos ESA. 2012. Methodological issues and advances in biological meta-analysis. *Evol Ecol.* 26:1253–1274.
- Noble DWA, Lagisz M, O’dea RE, Nakagawa S. 2017. Nonindependence and sensitivity analyses in ecological and evolutionary meta-analyses. *Mol Ecol.* 26:2410–2425.
- Réale D, Garant D, Humphries MM, Bergeron P, Careau V, Montiglio PO. 2010. Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philos Trans R Soc Lond B Biol Sci.* 365:4051–4063.
- Rosenthal MF, Gertler M, Hamilton AD, Prasad S, Andrade MC. 2017. Taxonomic bias in animal behaviour publications. *Anim Behav.* 127:83–89.
- Royauté R, Berdal MA, Garrison CR, Dochtermann NA. 2018. Paceless life? A meta-analysis of the pace-of-life syndrome hypothesis. *Behav Ecol Sociobiol.* 72:64.
- Sih A, Mathot KJ, Moirón M, Montiglio PO, Wolf M, Dingemanse NJ. 2015. Animal personality and state-behaviour feedbacks: a review and guide for empiricists. *Trends Ecol Evol.* 30:50–60.
- Stirling DG, Reale D, Roff DA. 2002. Selection, structure and the heritability of behaviour. *J Evol Biol.* 15:277–289.
- van Oers K, Sinn DL. 2013. Quantitative and molecular genetics of animal personality. In: Carere C., Maestripieri D., editors. *Animal personalities: behavior, physiology, and evolution*. Chicago: University of Chicago Press. p. 159–200.
- Viechtbauer W. 2010. Conducting meta-analyses in R with the metafor package. *J Stat Softw.* 36.
- Walsh B, Blows MW. 2009. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. *Annu Rev Ecol Evol Syst.* 40:41–59.
- West-Eberhard MJ. 2003. *Developmental plasticity and evolution*. New York: Oxford University Press.