Linking personality and cognition: a meta-analysis

Liam R. Dougherty1 and Lauren M. Guillette2,3

1Department of Evolution, Ecology and Behaviour, Institute of Integrative Biology, University of Liverpool, Liverpool L69 7RB, UK
2School of Biology, University of St Andrews, St Andrews, Fife KY16 9TH, UK
3Department of Psychology, University of Alberta, Edmonton, Canada AB T6G 2R3

In the past decade, several conceptual papers have linked variation in animal personality to variation in cognition, and recent years have seen a flood of empirical studies testing this link. However, these results have not been synthesized in a quantitative way. Here, we systematically search the literature and conduct a phylogenetically controlled meta-analysis of empirical papers that have tested the relationship between animal personality (exploration, boldness, activity, aggression and sociability) and cognition (initial learning/reversal speed, number of correct choices/errors after standard training). We find evidence for a small but significant relationship between variation in personality and variation in learning across species in the absolute scale; however, the direction of this relationship is highly variable and when both positive and negative effect sizes are considered, the average effect size does not differ significantly from zero. Importantly, this variation among studies is not explained by differences in personality or learning measure, or taxonomic grouping. Further, these results do not support current hypotheses suggesting that fast-explorers are fast-learners or that slow-explorers perform better on tests of reversal learning. Rather, we find evidence that bold animals are faster learners, but only when boldness is measured in response to a predator (or simulated predator) and not when boldness is measured by exposure to a novel object (or novel food). Further, although only a small sub-sample of papers reported results separately for males and females, sex explained a significant amount of variation in effect size. These results, therefore, suggest that, while personality and learning are indeed related across a range of species, the direction of this relationship is highly variable. Thus further empirical work is needed to determine whether there are important moderators of this relationship.

This article is part of the theme issue ‘Causes and consequences of individual differences in cognitive abilities’.

1. Introduction

In the past 15 years, research in behavioural ecology has shown that different behaviours of individual animals may be stable across time or contexts (animal personality sensu [1–3]). These different behaviours (also called personality traits), moreover, may not be independent from one another and, seemingly independent behaviours, measured using different tasks, could form suites of correlated traits (behavioural syndromes sensu [4–6]). Thus, the tide of studying the average behaviour of groups has ebbed, as researchers have realized the importance of quantifying the variation among individuals in a group [7]. Along with this upwelling of empirical papers on animal personality came a swell of conceptual, terminological and statistical papers (‘data-free’ papers, reviewed in [8]) linking personality to many aspects of ecological and evolutionary biology (e.g. sexual selection [9]; conservation [10]; ecology and evolution [11]; development [2]; evolutionary genomics [12]). Included in this swell are several conceptual papers linking animal personality to animal cognition [13–19].
A link between personality and cognition, albeit by different names, was first established by Pavlov in the early twentieth century during his work examining associative processes (i.e. conditioned reflexes) and digestive physiology [14,20–22]. Pavlov described four different ‘types’ of nervous systems based on how quickly dogs learned to form different types of associations [23]. For instance, the ‘Excitable type’ showed strong (and quick) excitatory conditioning (learning to make a response), but weak (and slow) inhibitory conditioning (learning to withhold making a response). The ‘Inhibited type’ was the opposite: showing strong and quick inhibitory conditioning, and weak and slow excitatory conditioning. Both the Excitable and Inhibited type also showed low flexibility—that is, alternating between excitatory and inhibitory conditioning.

The ‘Lively type’ showed rapid associative learning for both excitatory and inhibitory tasks and could make flexible conversions between the two. The last type, ‘Quiet’, formed slow but consistent associations and was less flexible, compared with the Lively type, when transitioning between the different conditioning types (excitatory and inhibitory; [13,14]). In two lectures: An attempt to understand the symptoms of hysteria physiologically (1932) and The conditioned reflex (1935; [22]), Pavlov connected the four types of nervous systems to individually distinct animal ‘temperaments’. For instance, the Excitable type display general behaviour that is ‘aggressive, animated and undisciplined’ [22, pp. 105], while the Lively and Quiet type behave ‘actively and lively’ and ‘inert...calm and unper- turbed’ [22, pp. 177], respectively. Lastly, the Inhibited type is ‘restless and constantly looking about or on the contrary, constantly stopping and remaining motionless . . . ’ [22, pp. 177]. Pavlov believed these four types of nervous systems were responsible for individually distinctive and fixed behavioural phenotypes (i.e. personalities) of different dogs [22].

The foremost goal of this paper is to assess if Pavlov was indeed correct by asking: is an animal’s personality related to its cognitive ability? Recent years have seen a flurry of empirical studies testing this question, in a range of species (e.g. mammals, [24]; fish, [25]; birds, [26]). However, these results have not yet been synthesized in a quantitative way. We address this using a meta-analytic approach. We systematically searched the literature for studies testing for the relationship between animal personality and cognition across individuals, finding estimates for 19 animal species, including mammals, birds, reptiles, fish and insects. We use data from papers examining at least one measure of personality. The proposed mechanism for this correlation is a risk–reward trade-off that underlies both cognition and personality, that is, the more a behaviour is expressed (e.g. more aggression, more boldness, fast-learner) the greater the risk (e.g. being predated, injury in contests, decision errors). Sih & Del Giudice [15] make a distinction between cognitive abilities and cognitive ‘style’, where cognitive style refers to ‘the way individuals acquire, process, store or act on information, independent of cognitive ability’ [15, pp. 2762]. And, while the distinction between ability and style is not usually discussed or addressed in papers examining cognition, the theoretical framework supplied by [15] is applied (see [16] for example of measuring cognitive style). A similar view, linking personality to cognition, holds that bold/explorative animals experience more of their

Importantly, the relationship (correlation) between personality and cognition can be either positive or negative, depending on how behaviours are coded. While the assignment of a direction to these behavioural measures is somewhat arbitrary (see Material and methods), the biological meaning is not; for example: a positive relationship between cognition and boldness (e.g. faster learners are bolder) is biologically and ecologically different from the converse (e.g. faster learners are less bold). However, another way to examine this relationship across species is to look at the absolute magnitude of the effect, irrespective of the sign (in other words by making all effect sizes positive). Such an approach may be needed if the sign of the relationship is not consistent across species [30,31]. In such a case, using the absolute values may allow us to detect a strong relationship that is masked when we examine the raw (positive and negative) effect sizes alone, and this result would be informative in that it suggests that there are underlying factors that strongly influence the direction of the relationship which we can try to uncover. In this study, we, therefore, quantify the strength of the relationship between personality and cognition both with and without considering the directionality of the effect sizes.

The secondary goal of this paper is to begin to address specific predictions regarding the direction of the relationship between personality and cognition. Although it has been argued elsewhere [16], making predictions about the direction of the relationship between personality and cognition will depend on many factors, including, but not limited to—the stimulus (e.g. tone, light, conspecific, odour), the response (e.g. making one versus withholding making one), and the outcome (positive or negative). A popular prediction, nonetheless, based on both conceptual [13,15,19,28,29] and early empirical work (e.g. [32,33]), is that fast-explorers are fast-learners and excel in stable environments, whereas slow-explorers are more flexible and, therefore, should be better at reversal learning compared with fast-explorers. In other words, the relationship between exploration and cognition may depend on the cognitive measure being used. Therefore, we predict a positive relationship between personality and learning speed for newly acquired tasks (e.g. fast-explorers are fast-learners) and a negative relationship between personality and reversal learning (e.g. slow-explorers are fast at reversal learning).
Figure 1. PRISMA diagram showing systematic search process. See electronic supplementary material for complete list of search terms used in different databases, and electronic supplementary material, table S1 for a list of relevant papers not included in the final analysis. For the articles excluded ‘Only one paradigm tested’ refers to papers where only personality, or cognition, but not both, were tested.

environment, more quickly, thus coming into contact with to-be-learned associations more readily than shy/less explorative individuals [13,15,16]. This view, therefore, suggests that personality constrains cognition. The same end can also be achieved by different means: animals that form associations more quickly may be able to then move through their environment more quickly (learning ability facilitates exploration, [27]). Despite these different proposed mechanisms, the outcome remains the same—a positive link between exploration and learning speed. In the current meta-analysis we ask if six different personality measures are related to learning in the same way.

In summary, in this study we ask three questions. First, is there a significant relationship between personality and learning, in either the absolute or raw scale? Second, is the strength or direction of this relationship influenced by additional factors, such as the personality measure or cognitive test used, or the sex of the subjects? Third, is there any evidence of publication bias against studies showing certain results (e.g. those that contradict prevailing theory)?

2. Material and methods

Our methods followed the PRISMA standards for reporting meta-analyses ([34–36]; see figure 1 for a diagram of the search results and study selection) as closely as possible.

(a) Search protocol

We used three methods to search the literature for relevant studies. First, keyword searches were performed using three databases on 17 October 2017 (Web of Science, PsychINFO and Scopus, see electronic supplementary material for complete list of search terms used for each database). Second, Web of Science was used to search for papers that had cited two influential papers in this area: a review on behavioural syndromes and cognition [15], and an opinion paper on cognition and personality [16]. After these searches, we excluded duplicate results, and then accessed the abstracts of 1776 papers and screened them for inclusion. Full texts of papers that were deemed relevant were read (n = 129). Finally, the full texts of three additional papers that were not located by the initial search were accessed because they were cited in the papers that were deemed relevant (final n = 132, figure 1).

(b) Criteria for inclusion

We had several criteria for inclusion of a study in our analysis (see electronic supplementary material, table S1 for a list of studies not included in the analysis, and the reasons for their exclusion). The main criterion was that each paper needed to include at least one measure of personality and one measure of cognition, which came from different tasks. For example, in a study examining boldness (as measured by latency to interact with a novel object) and learning speed (number of trials to reach criteria for a visual discrimination task), this criterion was violated if boldness was measured as latency to interact with the cognitive testing apparatus that was used to assess learning speed. Second, the paper needed to present statistical information so that an effect size could be calculated (though note that in several cases we contacted the authors of papers that did not present appropriate statistics in order to obtain such information; see below for more details).

(i) Personality measure

The relatively young field of animal personality faces several challenges when it comes to measuring personality, which is clearly reviewed in [29]. One challenge relates to defining personality traits, a second challenge relates to how these traits are measured (see [37] for discussion about failure to measure repeatability in traits and [38] for a meta-analysis of repeatability of personality traits). Here, we followed the definition of a personality trait from [29, pp. 476]: A specific aspect of a behavioural repertoire that can be quantified and that shows between-individual variation and within-individual consistency (such as boldness, aggression, activity). We included studies that report personality measures from one or several behavioural episodes. In other words, we did not specify that a personality measure needed to be shown to be repeatable in order to include it in our analysis, but instead defer to the original authors’ judgments. The terminology for the specific personality
traits used here is based on [11], sometimes referred to as the ‘Big Five’: boldness, exploration, activity, aggressiveness and sociability [6]. However, [11] explicitly addressed the limitation of this over-simplification of terminology and suggested that the five outlined traits be regarded as a working tool. Thus the working definitions we used are more in line with those used by [28], and consisted of the following categories: boldness—responses to novel objects, food and potential predators; exploration—responses to a novel environment or open field; social/aggression—reactions to con-specific presentations; activity—movement around a familiar environment (e.g. a home cage); and, exploration/boldness—combined reactions to novel environment and novel object tests (e.g. established composite scores for great tits sensu [39]). Note that in the analysis we distinguish between boldness in response to novel objects or food and boldness in response to predators, as preliminary analyses indicated that these were informative groupings. We use the term ‘personality measure’ rather than ‘personality trait’ in order to distinguish between these two types of boldness. In summary, the ‘behaviour measures’ variable consists of six categories: boldness in response to novel objects/food, boldness in response to predators, exploration/boldness, activity, exploration and social/aggression.

(ii) Cognitive measure and training type
We included studies that examined four different cognitive measures (learning speed, reversal learning speed, number of errors, number of correct responses)—which we grouped into two different training types: ‘trials to criterion’ and ‘standard training’. In the first type of study (trials to criterion), animals were trained until they reached pre-determined learning criteria: (1) for initial acquisition of a task (learning speed); or (2) during a subsequent phase when the initial reward contingencies (those in place during initial acquisition) were reversed (reversal learning). Animals trained to criteria are at the same level of asymptotic performance (e.g. in associative learning, the maximum associative value a conditioned stimulus can gain [40]). In the second type of study (standard training), animals were trained for a standard number of trials and the cognitive measures were: (3) the number of errors; or (4) the number of correct responses. In these latter two measures, it is unclear if or how much an animal has learned (i.e. where an individual’s performance falls on a learning curve that culminates, theoretically, in asymptotic learning). We, therefore, have separated these from the cases where animals are trained until they reach learning criteria. There is a dearth of studies that examine the link between cognitive abilities beyond information acquisition (i.e. information use, but see [41] for a test of generalization of previously learned rules in a pigeon and [42] for a test of performance on novel exemplars following initial acquisition). We did not include studies that tested motor learning or problem solving (extractive foraging task) as it is unclear which cognitive mechanism may underpin performance in these tasks (for an in-depth treatment of this topic, see [43–45]). Note that electronic supplementary material, table S3 indicates the Cognitive measure and Training type for all effect sizes in the meta-analysis (see reference [16, table 1] for overview of measurement of cognitive abilities).

(iii) Sex
We included both studies that tested for sex differences in behaviour and those that did not, with sex classified as ‘both’ when sex differences were not assessed. In one case, the sex of the subjects was not specified [46]; therefore, we classed this as ‘both’.

(c) Calculating effect sizes
In order to quantify the relationship between personality and learning, the experimental results first need to be converted into a standardized effect size. We used Pearson’s product moment correlation coefficient (r) as the measure of effect size, as the majority of studies in our sample measured both personality and learning on a continuous scale (though there were nine cases in which subjects were classified based on discrete groups based on a personality or cognition score). Here, r represents the magnitude of the association between one of several personality measures and some cognitive measure. Given that r can range from +1 to −1, we need to determine the sign of the relationship for each study. We classified correlations as either positive or negative depending on the following criteria. Positive effect sizes were assigned when individuals that had faster learning (or reversal) speeds, more correct choices, or fewer mistakes were also: more active, more explorative, bolder, more aggressive or more sociable. Negative effect sizes were assigned when individuals that had faster learning (or reversal) speeds, more correct choices or fewer mistakes were also: less active, less bold, less aggressive or less sociable. Note that individuals that were classed as ‘faster’ at learning took fewer trials to reach the learning criteria, but this is still classed as a positive effect size. The direction of effect was determined either by using the sign of test statistics presented in the papers, by using the descriptions given by the authors, or by examining the raw data.

If studies did not report r, it was computed from the available statistical information, or from additional information provided by the authors, using the procedures in [34]. See electronic supplementary material, table S2 for full details on the calculation of effect sizes when r was not reported. Only one paper (two effect sizes) reported r directly. Twenty-one effect sizes (from nine papers) were obtained by converting statistical data presented in the text. For the remaining 45 effect sizes, new calculations were made using descriptive statistics presented in the text (two papers) or raw data provided in the paper, in the accompanying electronic supplementary material, or by the authors (13 papers).

In 17 out of 25 studies we obtained more than one effect size. In all but one study [47] this was due to multiple tests being performed on the same sample of individuals. However, note that sample sizes often varied between tests from the same study, usually because some tests could not be performed using all individuals. When calculating the total number of individuals used in any study or data set (electronic supplementary material, table S4) we were, therefore, careful to avoid pseudo-replication by not counting any individual more than once. For all analyses, we used Fisher’s Z transform of the correlation coefficient (Zr), as this has better statistical properties when r approaches ±1 [34]. The associated variance for Zr (var Z) was calculated as 1/(r − 3) [48].

(d) Generating the phylogeny
Our sample included data from multiple species across several taxonomic classes, and therefore one potential confounding factor is similarity due to shared evolutionary history [34]. Modern meta-analytic methods allow the phylogenetic relatedness of species to be taken into account during the analysis [49]. However, as our sample includes a wide range of species, spanning several vertebrate classes (as well as a single invertebrate species), there is currently no single phylogeny available that incorporates every species included. We, therefore, constructed a supertree by manually combining multiple smaller trees from the literature. We used taxonomic groupings for species for which phylogenetic data were not available [49]. We obtained phylogenetic trees from several sources: for the relationship among birds we used [50,51]; for the relationship among fish we used [52]; for the relationship among mammals we used [53]; and for the relationship among vertebrates we used [54].

The supertree approach also means that obtaining accurate branch length data for the phylogeny is not possible. However, the phylogenetic branching pattern of the tree still contains important information on the relatedness between different taxa [36], and so we estimated branch lengths based on the total length of the tree [55]. Accordingly, we first assigned all branch lengths a
value of 1. The tree was then made ultrametric (all tips contemporaneous), and branch lengths estimated, using Grafen’s method [55], by means of the analysis of phylogenetics and evolution (APE) package v. 3.3 [56] in R v. 3.5. The final ultrametric tree used in the analysis is shown in figure 2.

(e) Statistical analysis

All analyses were performed using R v. 3.5 (R Core Development Team, 2018) and Metafor v. 1.9 [57]. Meta-analysis models were run using a Bayesian approach, using the package MCMCglmm v. 2.21 [49]. We first ran a multilevel meta-analysis model in order to estimate the mean effect size across all studies in the sample. We use the term ‘multilevel’ to refer to random-effects meta-analysis models (in traditional meta-analysis classification; see [34,48]) that include additional random factors in order to control for potential non-independence between effect sizes (following [58]). We included study, species and phylogenetic relatedness (using the phylogenetic tree shown above) as random factors in these models. Study was included as a random factor because we extracted more than one effect size from most studies (average of 2.64 effect sizes per study, range = 1–6). Species was included as a random factor because four species (Cavia porcellus, Passer domesticus, Poecile atricapillus and Taeniopygia guttata) were tested in more than one study. Phylogeny was included as a random factor as our sample included several species in the same genus/family. Removing any of these random factors did not significantly improve model fit, or influence the significance of any categorical factors in meta-regression models (see below); therefore, we included all the three random factors in all models.

All models were fitted using an inverse-Wishart prior for all fixed and random effects ($V = 1, \nu = 0.002$, [31,59]). All models were run for 3 million iterations, with a thinning interval of 2000 and a burn-in period of 2 million iterations. We present our results as mean posterior estimates of $r$ (back-converted from $Z$, after analysis), as well as the highest posterior density (HPD) interval (also referred to as the 95% credible interval). We consider an estimate to be significantly different from zero if the HPD credible interval does not overlap zero. We checked the convergence of all models by examining the Markov chain Monte Carlo (MCMC) time series; the number of iterations was sufficient to result in no trend for any of the models. We checked model mixing by checking the autocorrelation between the stored samples in the chain (representing the end of the MCMC run). Values for all models were less than 0.1, indicating good mixture. We ran all models three times using identical parameters, and used Gelman–Rubin diagnostics to check for convergence between the three runs [60]). These diagnostics produced a potential scale reduction factor point estimate of 1 or very close to 1, indicating convergence. We also re-ran the intercept-only model using a flat prior for the residuals and random effects ($V = 1 \times 10^{-18}, \nu = -2$), with the same number of iterations as all previous models. This model gave a very similar mean estimate to those using an inverse gamma prior, though the credible interval was significantly wider, and we do not present it here.

We assessed the amount of heterogeneity in effect sizes for the intercept-only model using the $I^2$ statistic [61]. This statistic estimates the percentage of overall variation in the sample that is due to heterogeneity between studies (or effect sizes in this case) compared with sampling error (variation within studies). The $I^2$ value is generally preferred over Cochran’s $Q$ test, as it gives an estimate of the degree of heterogeneity, rather than just a $p$ value, and is less affected by sample size. We present $I^2$ values associated with the overall model, and each of the three random factors, following [58]. We follow the recommendations of [61] in considering $I^2$ values of 25, 50 and 75% as low, moderate and high respectively, though heterogeneity in ecological and evolutionary meta-analyses is typically very high [62].
The first analysis was used to estimate both the magnitude and the direction of the relationship between cognition and personality. However, given that the sign of the effect was highly variable (see below), and there are not always clear predictions for which direction this relationship should take, we also wanted to estimate the absolute magnitude of the relationship between personality and cognition ($r$), irrespective of the sign. We did this by applying the folded normal distribution to the posterior mean estimate derived from the intercept-only model, in order to estimate the average effect size and credible intervals on the absolute scale (i.e. the ‘analyse and transform’ approach recommended by [30,31,37,63]).

We next examined the extent to which variation in effect size was related to five categorical moderator variables. These were personality measure, cognitive measure, taxonomic class, sex and training type (see ‘criteria for inclusion’ for category details).

We performed a series of meta-regression models, each of which included study, species, and phylogeny as random effects, and one of the five categorical fixed effects. Model fit was then determined using the deviance information criterion (DIC), which is a Bayesian equivalent of traditional information theoretic criteria. Lower values indicate a better fit, and a change in DIC of two or more (compared with the multilevel model without moderators) was considered to indicate a significant improvement in model fit [64]. In order to obtain mean effect size estimates for each factor level we also ran five mixed-effects models, each including only a single fixed effect, and with the intercept excluded. Again, we consider an estimate to be significantly different from zero if the HPD credible interval does not overlap zero. We also applied the folded normal distribution to the posterior mean estimates from these models in order to estimate the average magnitude ($r$) for each category of the five moderator variables. Finally, we calculated the amount of variance explained by the fixed factors (marginal $R^2$) using the method of [65].

We examined the dataset for two types of publication bias. First, we looked for evidence of bias against publishing studies with small or negative effect sizes, or with small sample sizes. To do this we tested for a relationship between effect size and variance using a rank correlation test [66] and a linear regression test [67]. However, these methods assume that effect sizes are independent, which does not apply to our dataset. Therefore, we used meta-analytic residuals rather than the raw effect sizes [58]. We also used the trim-and-fill method to test for asymmetry in the ‘funnel plot’ of residual effect size against sample variance. Asymmetry in the funnel plot is assumed to be indicative of publication bias against the ‘missing’ effect sizes on either side of the plot [68].

However, these methods assume that effect sizes are independent, which does not apply to our dataset. Therefore, we used meta-analytic residuals rather than the raw effect sizes [58]. We also used the trim-and-fill method to test for asymmetry in the ‘funnel plot’ of residual effect size against sample variance. Asymmetry in the funnel plot is assumed to be indicative of publication bias against the ‘missing’ effect sizes on either side of the plot [68], although there are other reasons for such asymmetry [58]. Second, we tested whether there is a relationship between effect size and the year the study was published, which may be indicative of publication bias. For example, the commonly observed negative relationship between effect size and year may be due to a greater bias against publishing studies of small effect in the early stages of the development of a new theory [34,69]. We examined this temporal trend by performing a meta-regression of the raw correlations, with year of publication added as a fixed factor and study, species and phylogeny as random factors.

**3. Results**

(a) Final dataset

The final dataset consisted of 25 studies and 66 effect sizes, testing 652 individuals in total. This included data for 19 species across a broad taxonomic range, including insects [70], fish [46,71–75], reptiles [76], birds [33,42,47,77–85] and mammals [86–90].

Figure 3. Funnel plot showing the relationship between sample size (inverse standard error; studies with larger sample sizes have larger values) and raw effect size ($Z_r$). The solid line shows the overall mean effect size estimate from a multilevel meta-analysis including all 66 effect sizes.

(b) Overall relationship

The overall mean effect size was not significantly different from zero ($r = 0.098$, HPD interval = $-0.074–0.281$, $N = 66$ effect sizes, 652 individuals). It can be seen from the funnel plot (figure 3) that the sample consists of an approximately equal number of positive and negative effect sizes. The overall heterogeneity of effect sizes ($I^2$) was moderate to high ($I^2 = 67.09\%$, HPD interval = $49.1–80.39\%$). It is, therefore, unlikely that this heterogeneity has arisen due to sampling error alone. The three random factors explained little of the heterogeneity in effect sizes (study $I^2 = 8.46\%$, HPD interval = $0.16–31.98\%$; species $I^2 = 5.21\%$, HPD interval = $0.12–17.52\%$; phylogeny $I^2 = 10.71\%$, HPD interval = $0.24–37.69\%$). The absolute mean effect size ($|r|$) was 0.268 (HPD interval = 0.179–0.368, significantly different from zero, $N = 66$ effect sizes, 652 individuals), which is considered medium to small (small effect size of 0.1, medium effect size of 0.3; [91]).

(c) Moderator variables

Given the high heterogeneity in effect sizes, we searched for potential moderators of this heterogeneity using a model selection approach. The variance explained by the fixed factors was low for all models, and sex was the only categorical factor that significantly improved model DIC (electronic supplementary material, table S5). Accordingly, there is a significantly positive relationship between learning and personality when males were tested ($r = 0.511$, HPD interval = $0.239–0.75$, $N = 4$ effect sizes, 98 individuals; figure 4), but not when females were tested ($r = 0.012$, HPD interval = $-0.298–0.308$, $N = 8$ effect sizes, 103 individuals), or when the sexes were not considered separately ($r = 0.064$, HPD interval = $-0.098–0.251$, $N = 54$ effect sizes, 511 individuals). However, the positive effect seen in males is due to only four effect sizes of large effect. When examining the personality measures category separately, there was a marginally significant positive relationship between learning and boldness in response to predators, with a mean $r = 0.363$, HPD interval = $-0.016–0.641$, $N = 5$ effect sizes, 98 individuals; figure 4). All other categories tested had mean effect size estimates that did not significantly differ from zero (figure 4). The absolute average effect size ($|r|$) across all behavioural measure categories was generally between 0.2 and 0.4, with the exception of effect sizes considering males ($|r| = 0.549$, HPD interval = 0.298–0.744), females ($|r| = 0.44$, HPD interval = 0.242–0.713) and...
fish ($j = 0.451$, HPD interval $= 0.261–0.725$, $N = 11$ effect sizes, 154 individuals; figure 5).

(d) Publication bias

There was no significant relationship between residual effect size ($Z_r$) and study precision (Egger’s test: $t_{64} = −0.473$, $p = 0.64$; Begg–Mazumdar test: Kendall’s $τ = 0.033$, $p = 0.7$). Further, trim-and-fill analysis did not detect missing effect sizes on either side of the funnel plot. There was no significant relationship between raw effect size ($Z_r$) and year (meta-regression, fixed effect of year, $β = −0.024$, HPD interval $= −0.054–0.007$; electronic supplementary material, figure S1).

4. Discussion

Our analysis provides the first quantitative test of the relationship between personality and cognition in animals, using a sample of 25 studies and 19 species. We find evidence for a small but significant relationship between variation in personality and variation in learning across species in the absolute scale (i.e. irrespective of the sign of the effect sizes). However, the direction of this relationship is highly variable, so that the average effect size for the raw data is not significantly different from zero. This means that our sample includes an approximately equal number of studies showing a positive relationship between personality and cognition (e.g. animals that were more bold, aggressive, explorative, active and social were quicker to learn, or had fewer errors, or more correct responses after a standard amount of training) as showing a negative relationship (animals that were more bold, aggressive, explorative, active and social were slower to learn, had more errors, or fewer correct responses after a standard amount of training). Further, taking into account the type of personality measure or cognitive measure did not significantly explain the variation in the direction of this relationship seen across studies. Taken together, these results show that, while personality and learning covary significantly across the studies sampled here, there is currently no evidence for a consistent positive or negative relationship across species.
Given the large amount of variation in effect sizes seen in our sample, we included several categorical moderator variables in our analysis in order to examine whether they could significantly explain some of the variation in the size or direction of the relationship between personality and cognition. We had two key predictions regarding how these variables might influence this relationship. Our first prediction was that the relationship between personality and cognition should depend on the type of learning test used to measure cognition: with a positive relationship predicted between personality and initial learning speed, and a negative relationship predicted between personality and reversal speed. However, this prediction was not supported: cognitive measure did not significantly influence the direction of the relationship between personality and cognition. This finding is in direct contrast with conceptual work that suggests ‘fast’ personality types are ‘fast’ and ‘inflexible’ learners, with ‘inflexible’ meaning animals that persevere in previously rewarded patterns of behaviour (early empirical paper: [19]) or fail to produce new, correct behaviour when the rules of a task or the environment change or are altered (conceptual papers: [15,92]).

Our second prediction was that certain personality measures, notably exploration, are more likely to covary with cognition than others. However, this was not seen to be the case, with personality measure explaining little of the heterogeneity in effect sizes seen across species. However, we did find evidence for a marginally significant positive relationship between cognition and boldness in response to predators: animals that are bolder are able to learn new associations (and reverse previously learned associations) more quickly, and show more correct responses (and fewer errors) during standard training, compared with animals that are less bold. However, it should also be noted that this category consists of only five effect sizes from three studies, and so should be investigated further before any strong conclusions are made. Nevertheless, this result was in contrast to the other personality measures (activity, exploration, sociality and aggression), which all have mean effect sizes that are not significantly different from zero (including boldness when measured as a response to novel objects or food), and it is not clear why boldness in response to predation shows a significant directional relationship with cognition while the others do not. It is worth stressing here that we do not assume a causal direction for this relationship—for example, it is equally likely that being a fast learner could lead individuals to be bolder.

The only categorical factor that explained a significant amount of the variation in effect sizes in our sample was the sex of the subject. For the directional data, the relationship was significantly positive when only males were tested, whereas the relationship for females and both sexes combined
did not significantly differ from zero. Further, the absolute size of the relationship between personality and cognition was more positive when males or females were tested separately, compared with when individuals of both sexes were combined. This result is somewhat surprising, given that there have been few studies examining sex differences in the relationship between personality and cognition, and indeed only a single study in our sample tested for this relationship in males and females separately [81]. For this reason, and the fact that this effect is primarily driven by the presence of a relatively few effect sizes of large effect (four and eight effect sizes for males only and females only, respectively), we interpret this result cautiously. Nevertheless, we suggest that this pattern merits further investigation, and that researchers should test for sex differences, including interactions between sex and personality, in the relationship between personality and cognition before data from males and females are combined, and report this in the methods or results sections even when there is no significant difference. Sex differences in cognitive abilities has long been a well-studied area in human psychology [93] and is beginning to receive attention in studies of animal cognition (e.g. [94,95,96]).

Importantly, the majority of the variation in effect size and direction in our sample remains unexplained, with effect size not influenced by differences in personality measure, cognitive measure or phylogenetic history across studies. There are two potential explanations for this, either: the relationship between personality and cognition does not have a consistent ‘direction’, in which case we need to adapt current theory in order to explain this; or there are additional moderating factors that we have not identified that strongly influence the direction of the relationship. For example, given the limited size of our sample we did not test the effect of any ecological or life-history factors that may influence this relationship (e.g. sociality, breeding system, habitat type). Further, many of these studies tested a relatively small number of individuals; the average sample size across all studies was 26.08 (s.d. = 13.89), with eight studies testing fewer than 20 individuals. This means that many of the trait categories we examined consisted of a very small number of individuals (e.g. 45 individual insects and 57 individual reptiles). Therefore, we suggest that more empirical tests are needed to investigate these potential explanations, using larger sample sizes if possible. This is still a relatively young field, as exemplified by the fact that 19 of the 25 studies included in our analysis were published in the past five years, and there is much we still do not know. Nevertheless, other meta-analyses have shown that personality is related to an individual’s intrinsic state (i.e. body mass, size, metabolic rate and hormone levels; [37]) and has fitness consequences (e.g. reproductive success and survival; [28]). Taken together with the current results, this suggests that personality is a measure worth examining in the future.

In conclusion, our results show that Pavlov was correct: animal personality and cognition are related. However, our analysis also revealed high among-study heterogeneity in the direction of this relationship. This means that knowing the personality of an animal (where an individual’s behavioural scores fall along a continuum ranging from inactive to active, for example) does not consistently allow you to predict how quickly that animal will learn. Further, we failed to find support for several key hypotheses regarding the relationship between personality and cognition, and we hope that these hypotheses will be re-assessed accordingly. Specifically, researchers may need to abandon the primary assumption that fast-explorers should be fast-learners, while slow-explorers should be better at reversal learning tasks. Finally, further work is needed in order to identify whether there are other factors that influence the direction of the relationship between cognition and personality. In the light of these results, we have several recommendations. First, we urge researchers undertaking future work to test for sex differences and interactions between sex, personality and cognitive measures. Secondly, we suggest researchers measure both personality and cognition across several different time points, or in several different contexts in the same individuals (see [16] for details, and Cauchoox et al., this issue [97]). Lastly, our hope is that this meta-analysis stimulates empirical work where formulation of study-specific predictions should take into account not only the evolutionary pressures that have shaped different species’ cognitive abilities, but also the different developmental histories among discrete populations of the same species (e.g. pond snails, Lymnaea stagnalis [98]; Dalesman, current issue [99]); sticklebacks, Gasterosteus aculeatus, [100,101]) along with the nature of the cognitive testing paradigm (the stimuli, the behavioural response and the outcomes [16]).

Data accessibility. The dataset supporting this article has been uploaded as part of the electronic supplementary material.

Authors’ contributions. L.M.G. conceived the idea. L.M.G. and L.R.D. designed and collected the data and wrote the paper. L.R.D. analysed the data. All authors give final approval for this publication.

Competing interests. We have no competing interests.

Funding. This study was funded by LMG Biotechnology and Biological Sciences Research Council Annivary Future Leader Fellowship (BB/M013944/1) and LRD Association for the Study of Animal Behaviour Research Grant.

Acknowledgements. We thank Jo Madden for organizing an intellectually stimulating (and fun) workshop on variation in cognitive abilities. We gratefully thank the following authors for their correspondence and in many cases for providing additional data and information: Kim Mathot, Rachel Miller, Julie Gibelli, Julie Monard-Ferron, Vera Brust, Anja Guenther, Alon Shamir, Sergey Budaev, Patrizia d’Ettorre, Claudio Carese, Sara Bebus, Christian Nawroth, Alan McElligot, Laura Marina Bondi, Alice Exnerová, Cairest DePasquale, Mary Olmstead, Marc Naguib and Lee Alan Dugatkin. We thank Daniel Noble and Michael Morrissy for statistical advice and guidance, and Marco Del Giudice and two anonymous reviewers for thoughtful and useful comments on an earlier version of the manuscript.

References


