

# Criteria for acceptable studies of animal personality and behavioural syndromes

Studies involving animal personalities and behavioural syndromes are being carried out in large numbers and submitted to behavioural journals like *Ethology*. It therefore seems useful to consider the most important criteria for acceptable publications in this area of research. Much of this information is already available in various well-cited methodological papers (see below), but as reviewers and editors of many of these submissions we felt that it might be useful to provide a brief summary of some key issues. Our aim here is to assist authors in avoiding common problems and help them to produce strong publishable papers in personality research. These criteria are also intended to serve as up-to-date guidelines for reviewers and editors of *Ethology* and other similar peer-reviewed journals.

## 1 | DEFINITIONS AND STUDY DESIGNS

There is consensus that “animal personality” refers to the repeatable part of an individual's behaviour (Dingemanse, Kazem, Réale, & Wright, 2010; Réale, Reader, Sol, McDougall, & Dingemanse, 2007), with “variation in personality” referring to among-individual differences in average behaviour. Personality studies should therefore involve repeated measures designs. This enables statistical assessment of an individual's average behaviour (Dingemanse & Dochtermann, 2013), which may then in turn be appropriately linked to the scientific concept of interest.

Studies that do not feature repeated measures of the behaviour(s) concerned are not formally studies of animal personality, and should perhaps not be presented as such (Niemelä & Dingemanse, 2018a). Such studies may still usefully link their work to animal personality research provided that the authors highlight that their conclusions are conditional upon the assumption that the single measurement of behaviour per individual is only a proxy of the true behavioural mean (or “personality”) of each individual (Niemelä & Dingemanse, 2018b). Unfortunately, this assumption is unlikely to hold in most instances, because behavioural repeatability is on average only about 0.4 (Bell, Hankison, & Laskowski, 2009; Holtmann, Lagisz, & Nakagawa, 2017), implying that a behaviour measured only once largely reflects the within (60%) rather than among (40%) individual component (Dingemanse, Dochtermann, & Nakagawa, 2012). For this reason, the argument “we have previously shown that this behaviour is individually repeatable and thus we can use a single measure per individual as a proxy for its personality” is neither reasonable nor logical, because the single measure per individual will

largely reflect the “unrepeatable” portion of individual phenotypes. A suitable degree of caution is therefore required in all such cases. We illustrate this point with a thought experiment, focusing on circulating levels of corticosterone and its apparent effects on fitness in a population of birds (because hormonal traits have particularly low repeatability ( $R = 0.1\text{--}0.2$ ); Holtmann et al., 2017). We might detect a strong negative relationship between fitness and individual levels of corticosterone when measured only once, and from that we could conclude the trait is under strong directional selection for reduced levels of corticosterone. This would obviously be a false conclusion if the repeatable part of corticosterone did not actually predict fitness. Instead, the real explanation could well be that harsh environmental conditions reduced reproductive success whilst simultaneously also temporarily elevating individual levels of corticosterone (this problem is well-known in evolutionary biology where it is standard practice to control for such environmental confounds when measuring selection; Rausher, 1992; Stinchcombe et al., 2002). The potential for these types of biasing effects of environmental confounds is thus particularly high for traits with low repeatability, which includes many of the behaviours in studies exploring these questions with regard to animal personality.

An important associated issue here is that many behaviours exhibit plasticity in response to environmental factors that change relatively slowly (compared to the lifetime of the species), creating strong temporal (Allegue et al., 2017; Mitchell, Dujon, Beckmann, & Biro, 2019) or spatial (Niemelä & Dingemanse, 2017) autocorrelations. Individuals may therefore be repeatable only because of being assessed in the same repeatable environmental conditions to which they are responding plastically (Dingemanse, Kazem, et al., 2010; Martin & Réale, 2008) and not because of any genuine among-individual variation caused by genetics or developmental effects (Dochtermann, Schwab, Berdal, Dalos, & Royaute, 2019; Dochtermann, Schwab, & Sih, 2015; Stamps & Frankenhuis, 2016). This phenomenon is called “pseudo-repeatability” or “pseudo-personality” (Dingemanse & Dochtermann, 2013; Westneat, Hatch, Wetzel, & Ensminger, 2011) and has been firmly documented by meta-analyses showing that repeatability values decrease with increasing length of inter-test intervals (Bell et al., 2009). Research primarily centring on the ecological or evolutionary causes or consequences of “genuine” repeatable differences (as defined above) should therefore avoid pseudo-personality. This can be achieved by using study designs allowing statistical control for biasing effects of environmental spatiotemporal autocorrelations (Allegue et al., 2017;

Araya-Ajoy, Mathot, & Dingemanse, 2015; Mitchell et al., 2019). Generally, a demonstration of animal personality is therefore most convincing when the inter-test interval is relatively long compared to the lifespan of the species (Allegue et al., 2017). For instance, it may not be reasonable to collect repeated measures a few days apart and then draw conclusions regarding how animal personality affects aspects of fitness or life history across weeks or months. Personality studies must therefore acknowledge that animals are both plastic and repeatable, and aim to separate these two aspects of behaviour (Dingemanse, Kazem, et al., 2010).

Some studies define “animal personality” as pertaining to a particular class of behaviours regardless of repeated measures, often (inappropriately) citing Réale et al. (2007) in support. For example, behaviours related to coping with captive or novel conditions, risk-taking, boldness or aggression are often considered “personality traits.” However, such approaches do not easily integrate into the sub-field that has emerged within behavioural ecology that involves the development of adaptive explanations for repeatable individual variation (“personality”) in *any* behaviour (Dall, Houston, & McNamara, 2004). Furthermore, some studies refer to “personality traits” and define them as “behavioural traits that are repeatable.” This labelling is arguably not helpful because *all* evolved behaviours are repeatable provided we collect enough data (Falconer & Mackay, 1996), and thus it effectively replaces one label (“behaviour”) with another (“personality”). Defining variation in “animal personality” as among-individual variation in average behaviour across repeated observations (as detailed above) provides a much more useful terminology that does not suffer from such issues.

Many of the points above also apply to behavioural syndromes (for a recent discussion, see Niemelä & Dingemanse, 2018a). In short, a “behavioural syndrome” is the among-individual correlation in repeated measures data across multiple behaviours (Dingemanse, Kazem, et al., 2010). It essentially represents the correlation between individual-mean values across suites of traits. As above, this requires repeated measures data for each behaviour (Dingemanse & Dochtermann, 2013). It also requires that the different behaviours are measured across separate assays; otherwise, the correlations may merely reflect different mutually inclusive or exclusive types of behaviour occurring during the same observation. Following our argument above, a correlation between two behaviours each measured only once probably reflects mostly the within-individual rather than the among-individual pattern of covariance (Brommer, 2013; Dingemanse et al., 2012). This is not a problem provided that among- and within-individual correlations do not differ (Brommer & Class, 2017; Dingemanse & Dochtermann, 2013). Unfortunately, meta-analyses often show that this assumption does not hold (Dochtermann, 2011; Niemelä & Dingemanse, 2018a). Again, studies involving single measurements per behaviour per individual may be usefully presented in the context of “behavioural syndromes,” provided that the caveats (assumptions) involved are clearly laid out and any conclusions in this regard are suitably toned down.

## 2 | STATISTICAL ANALYSES AND REPORTING RESULTS

Given the arguments above, it is clear that studies on animal personality and behavioural syndromes should normally report the among- and within-individual variance underpinning the behavioural repeatability reported in the paper (Dingemanse & Dochtermann, 2013) and coefficients of variation for each hierarchical level (Dochtermann & Royauté, 2019). The best approach currently available is to run a univariate mixed-effects model, where the explicit recommendation is to facilitate understanding of these results and any future meta-analyses by presenting all fixed and random effect parameter estimates along with their uncertainties (e.g., 95% credible intervals), information on the amount of variance that each component explains, and estimates of (adjusted) repeatability (Nakagawa & Schielzeth, 2010, 2013). Some studies instead choose non-parametric statistics, for example because distributional assumptions (e.g., of normality) were not met. It is important to understand that this precludes estimating key parametric parameters (e.g., repeatability, heritability and Pearson's correlations) that are firmly embedded in evolutionary theory (Dochtermann & Roff, 2010) and thereby diminishes the value of the study (Dingemanse et al., 2012). A suitable compromise may then be to run parametric analyses (e.g., mixed-effects models) and use bootstrapping or randomization procedures to derive appropriate null distributions for statistical testing. Fortunately, recent research suggests that mixed-effects models are, in fact, extremely robust to violations of distributional assumptions (Schielzeth et al., 2020).

Another common approach is to record multiple measures of behaviour within the same observation period or test assay, which are then aggregated using principal component analysis (PCA) into a single component or latent variable to represent an overall behavioural score. We have often used this approach ourselves in the past as a means of summarizing data. For example, one could estimate the heritability of a PCA component (e.g., Dingemanse, Barber, & Dochtermann, 2020) and link it to further aspects under study. This approach, however, becomes problematic when PCA components are estimated from data sets containing repeated measures since this represents a major violation (Budaev, 2010) and comes with (unreasonable) further assumptions (e.g., that correlation structures do not differ among versus within individuals, see above). Extreme caution should therefore be taken when applying such data-reduction approaches. A more appropriate approach would arguably be to report the correlation matrix between the different behavioural measures recorded within the same assay (within and among individuals) and then pragmatically take forward one representative behavioural measure that is normally distributed and covaries (i.e., loads) most strongly with the overall PCA component or latent variable estimated (Araya-Ajoy & Dingemanse, 2014; for other alternative approaches, see Dochtermann & Nelson, 2014). Either way, following the arguments above, a PCA cannot be used to imply that there is a “behavioural syndrome” or “animal personality structure”

if the study does not feature repeated measures and separate behavioural assays.

To our current knowledge, behavioural syndromes are best estimated by fitting multivariate mixed-effects (GLMM) models to repeated measures data across multiple traits (Dingemanse & Dochtermann, 2013). In such cases, it is important to report the variances, covariances and correlations at all levels (within and among individuals) either in the main text or supplementary materials. This enables meta-analyses of the entire matrix structure (e.g., Dochtermann & Dingemanse, 2013; Royauté, Hedrick, & Dochtermann, 2020) and comparisons across biological levels (Berdal & Dochtermann, 2019; Niemelä & Dingemanse, 2018a). As above, it is the among-individual correlation that best represents the behavioural syndrome in these instances, because it also takes into account additional (otherwise biasing) sources of variance and covariance within such data (e.g., Careau & Wilson, 2017; Downs & Dochtermann, 2014). Alternative statistical measures are sometimes used to estimate behavioural syndrome structures, which may be problematic as they can lead to biased estimates (Downs & Dochtermann, 2014). Unfortunately, multivariate mixed-effects models are extremely data hungry, implying that they can only be applied to large data sets. This requirement may inherently lead to publication bias towards species where such data can readily be collected. Such issues are worrying, but may partly be overcome by trading-off investments in other research conducted within the same laboratory, collaborative team work across laboratories or other smart solutions, and many of these choices are readily available to most principal investigators in this field (Niemelä & Dingemanse, 2018b).

An important problem in behavioural syndrome research is how to deal with large individual-level correlation or variance-covariance matrices, whether they are derived from multivariate (GLMM) models or not. Unfortunately, using PCA to summarize the major axes of variation in these correlation matrices does not allow assessments of how well such multivariate data fit specific alternative biological hypotheses. A more suitable statistical tool in this case is structural equation modelling (SEM) (Grace, 2006) or similar multivariate methods such as factor analysis or path analysis (Martin et al., 2019), because particular *a priori* hypotheses for specific covariance structures (e.g., based on the available literature) can then be tested and formally compared to find the most likely fit (Dingemanse, Dochtermann, & Wright, 2010; Dochtermann & Jenkins, 2007).

One further major issue is how measures of animal personality, in terms of differences in the mean behaviour per individual, are statistically linked with other physiological, ecological and evolutionary processes of interest. One common approach has been to first calculate the mean behaviour per individual and then statistically compare it in a separate model to some other biological variable, ecological measure or fitness, etc. This often involves doing statistics on what are essentially statistical estimates (i.e., "stats-on-stats"), and thereby ignores the error variance around both sets of measures. In order to avoid biased or misleading results, whenever possible the required estimates should be derived from a single

(multivariate) mixed model (Hadfield, Wilson, Garant, Sheldon, & Kruuk, 2010; Housley & Wilson, 2017). Unfortunately, these types of complex statistical models cannot always be applied to every data set, and this simply needs to be openly acknowledged in the paper. Suitable alternatives may be available, involving estimating the posterior distribution of any parameter and then taking this entire distribution (rather than just the point estimates) forward into subsequent analyses. For example, an increasingly popular approach is to determine the relative fit of alternative SEMs based on the entire posterior distribution (rather than the point estimates) of among individual (Araya-Ajoy & Dingemanse, 2014) or genetic correlation matrices (e.g., Dingemanse, Barber, et al., 2020; Dochtermann & Dingemanse, 2013). Such solutions are welcome, but can introduce their own biases and are therefore best used in combination with simulations that demonstrate their utility for the particular data set at hand (e.g., Araya-Ajoy & Dingemanse, 2017; Dingemanse, Moiron, Araya-Ajoy, Mouchet, & Abbey-Lee, 2020; Dochtermann & Dingemanse, 2013). Again, simpler pragmatic solutions are acceptable, but authors should always be clear about their different options, how any issues related to stats-on-stats were dealt with, and how those decisions might affect their biological conclusions.

As with all issues of study design and statistical analyses, it is important to note that research methods are constantly evolving (examples in this subject area include: Araya-Ajoy et al., 2015; Briffa, Rundle, & Fryer, 2008; Cleasby, Nakagawa, & Schielzeth, 2015; Dochtermann & Royauté, 2019; Mitchell et al., 2019). This summary is therefore the best current advice that we are aware of given our own reading and expertise. We ourselves have made many, if not all, of the aforementioned "mistakes" in our past publications, but we have always tried not to repeat them once a superior approach has been recommended. The point here is that researchers, and journal reviewers and editors, should stay up-to-date with currently acceptable standards within these types of fast-moving areas of research. However, we should all also review new methods critically (e.g., Class, Dingemanse, Araya-Ajoy, & Brommer, 2017; Cleasby et al., 2015; Dingemanse et al., 2012), and it is important that reviewers and editors alike are appropriately informed about study design and advanced statistical methods to be able to properly judge the merits of statistical approaches proposed or used in papers sent out for review (Bolker et al., 2009). Fortunately, this laudable aim is facilitated by accessible educational software packages, such as the R-package SQuID ("Statistical Quantification of Individual Differences"; Allegue et al., 2017), and by various "How-to" recipe papers (Dingemanse & Dochtermann, 2013; Wilson et al., 2010). However, an important message here is that we must guard against dogmatism and avoid paper rejection just because authors fail to use a popular method when they might have quite reasonable and defensible reasons for not doing so.

We would like to end by noting that although we have focussed primarily on issues concerning study design and statistical methods, our real motivation here is to improve the biological insights gained from investigating the evolutionary ecology of animal personality and behavioural syndromes. Authors submitting their work

for publication in *Ethology* and other behavioural ecological journals should therefore take note of one further requirement: the scientific question should clearly connect animal personality or behavioural syndromes to the ecology and/or evolution of the species involved. Studies that merely describe animal personality or behavioural syndromes are important, particularly as part of newly established study systems, but without these connections we would argue that they are not suitable for journals such as *Ethology*. Future studies should attempt to advance this field by publishing convincing tests of assumptions and predictions of adaptive theories in animal behaviour (Dall & Griffith, 2014; Dingemanse & Wolf, 2010). This editorial should therefore not be used to fuel criticisms of the field (Beekman & Jordan, 2017), but rather be seen as encouraging the continued submission of many high-quality studies in this valuable and flourishing area of behavioural research.

## ACKNOWLEDGEMENTS

We thank members of the Dingemanse laboratory and other participants of their daily digital coffee break for further feedback and discussion. We also thank members of SQuID for continued discussion of this topic over the years, and Denis Réale, Raphael Royauté and Tom Tregenza for comments on a previous draft of this Editorial. N.J.D. was supported by the German Science Foundation (grant no. DI 1694/1-1) and JW by the Research Council of Norway (SFF-III 223257/F50).

Niels J. Dingemanse<sup>1</sup>

Jonathan Wright<sup>2</sup>

<sup>1</sup>Behavioural Ecology, Department of Biology, Ludwig-Maximilians University of Munich, Planegg-Martinsried, Germany

<sup>2</sup>Department of Biology, Center for Biodiversity Dynamics, Norwegian University of Science and Technology (NTNU),

Trondheim, Norway

Email: jonathan.wright@bio.ntnu.no

## REFERENCES

- Allegue, H., Araya-Ajoy, Y. G., Dingemanse, N. J., Dochtermann, N. A., Garamszegi, L. Z., Nakagawa, S., ... Westneat, D. F. (2017). Statistical Quantification of Individual Differences (SQuID): An educational and statistical tool for understanding multilevel phenotypic data in linear mixed models. *Methods in Ecology and Evolution*, 8, 257–267. <https://doi.org/10.1111/2041-210X.12659>
- Araya-Ajoy, Y. G., & Dingemanse, N. J. (2014). Characterizing behavioural 'characters': A conceptual and statistical framework. *Proceedings of the Royal Society of London Series B*, 281, 20132645.
- Araya-Ajoy, Y. G., & Dingemanse, N. J. (2017). Repeatability, heritability, and age-dependence of seasonal plasticity in aggressiveness in a wild passerine bird. *Journal of Animal Ecology*, 86, 227–238.
- Araya-Ajoy, Y. G., Mathot, K. J., & Dingemanse, N. J. (2015). An approach to estimate short-term, long-term and reaction norm repeatability. *Methods in Ecology and Evolution*, 6, 1462–1473.
- Beekman, M., & Jordan, L. A. (2017). Does the field of animal personality provide any new insights for behavioural ecology? *Behavioral Ecology*, 28, 617–623.
- Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: A meta-analysis. *Animal Behaviour*, 77, 771–783.
- Berdal, M., & Dochtermann, N. A. (2019). Adaptive alignment of plasticity with genetic variation and selection. *Journal of Heredity*, 110, 514–521.
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology and Evolution*, 24, 127–135.
- Briffa, M., Rundle, S. D., & Fryer, A. (2008). Comparing the strength of behavioural plasticity and consistency across situations: Animal personalities in the hermit crab *Pagurus bernhardus*. *Proceedings of the Royal Society of London Series B*, 275, 1305–1311.
- Brommer, J. E. (2013). On between-individual and residual (co)variances in the study of animal personality: Are you willing to take the phenotypic gambit? *Behavioral Ecology and Sociobiology*, 67, 1027–1032.
- Brommer, J. E., & Class, B. (2017). Phenotypic correlations capture between-individual correlations underlying behavioral syndromes. *Behavioral Ecology and Sociobiology*, 71, 8.
- Budaev, S. V. (2010). Using principal components and factor analysis in animal behaviour research: Caveats and guidelines. *Ethology*, 116, 472–480.
- Careau, V., & Wilson, R. S. (2017). Of uberfleas and krakens: Detecting trade-offs using mixed models. *Integrative and Comparative Biology*, 57, 362–371.
- Class, B., Dingemanse, N. J., Araya-Ajoy, Y. G., & Brommer, J. E. (2017). A statistical methodology for estimating assortative mating for phenotypic traits that are plastic or measured with error. *Methods in Ecology and Evolution*, 8, 1910–1919.
- Cleasby, I. R., Nakagawa, S., & Schielzeth, H. (2015). Quantifying the predictability of behaviour: Statistical approaches for the study of between-individual variation in the within-individual variance. *Methods in Ecology and Evolution*, 6, 27–37.
- Dall, S. R. X., & Griffith, S. C. (2014). An empiricist guide to animal personality variation in ecology and evolution. *Frontiers in Ecology and Evolution*, 2, 3.
- Dall, S. R. X., Houston, A. I., & McNamara, J. M. (2004). The behavioural ecology of personality: Consistent individual differences from an adaptive perspective. *Ecology Letters*, 7, 734–739.
- Dingemanse, N. J., Barber, I., & Dochtermann, N. A. (2020). Non-consumptive effects of predation: Does perceived risk strengthen the genetic integration of behaviour and morphology in stickleback? *Ecology Letters*, 23, 107–118.
- Dingemanse, N. J., & Dochtermann, N. A. (2013). Quantifying individual variation in behaviour: Mixed-effect modelling approaches. *Journal of Animal Ecology*, 82, 39–54.
- Dingemanse, N. J., Dochtermann, N. A., & Nakagawa, S. (2012). Defining behavioural syndromes and the role of "syndrome deviation" to study its evolution. *Behavioral Ecology and Sociobiology*, 66, 1543–1548.
- Dingemanse, N. J., Dochtermann, N. A., & Wright, J. (2010). A method for exploring the structure of behavioural syndromes to allow formal comparison within and between datasets. *Animal Behaviour*, 79, 439–450.
- Dingemanse, N. J., Kazem, A. J. N., Réale, D., & Wright, J. (2010). Behavioural reaction norms: Animal personality meets individual plasticity. *Trends in Ecology & Evolution*, 25, 81–89.
- Dingemanse, N. J., Moiron, M., Araya-Ajoy, Y. G., Mouchet, A., & Abbey-Lee, R. N. (2020). Individual variation in age-dependent reproduction: Fast explorers live fast but senesce young? *Journal of Animal Ecology*, 89, 601–613.
- Dingemanse, N. J., & Wolf, M. (2010). A review of recent models for adaptive personality differences. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 365, 3947–3958.
- Dochtermann, N. A. (2011). Testing Cheverud's conjecture for behavioral correlations and behavioral syndromes. *Evolution*, 65, 1814–1820.



- Dochtermann, N. A., & Dingemanse, N. J. (2013). Behavioral syndromes as evolutionary constraints. *Behavioral Ecology*, 24, 806–811.
- Dochtermann, N. A., & Jenkins, S. H. (2007). Behavioural syndromes in Merriam's kangaroo rats (*Dipodomys merriami*): A test of competing hypotheses. *Proceedings of the Royal Society of London Series B*, 274, 2343–2349.
- Dochtermann, N. A., & Nelson, A. B. (2014). Multiple facets of exploratory behavior in house crickets (*Acheta domesticus*): Split personalities or simply different behaviors? *Ethology*, 120, 1110–1117.
- Dochtermann, N. A., & Roff, D. A. (2010). Applying a quantitative genetics framework to behavioural syndrome research. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365, 4013–4020.
- Dochtermann, N. A., & Royauté, R. (2019). The mean matters: Going beyond repeatability to interpret behavioural variation. *Animal Behaviour*, 153, 147–150.
- Dochtermann, N. A., Schwab, T., Berdal, M. A., Dalos, J., & Royauté, R. (2019). The heritability of behavior: A meta-analysis. *Journal of Heredity*, 110, 403–410.
- Dochtermann, N. A., Schwab, T., & Sih, A. (2015). The contribution of additive genetic variation to personality variation: Heritability of personality. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 282, 20142201.
- Downs, C. J., & Dochtermann, N. A. (2014). Testing hypotheses in eco-immunology using mixed models: disentangling hierarchical correlations. *Integrative and Comparative Biology*, 54, 407–418.
- Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics*. New York, NY: Longman.
- Grace, J. B. (2006). *Structural equation modeling and natural systems*. Cambridge, UK: Cambridge University Press.
- Hadfield, J. D., Wilson, A. J., Garant, D., Sheldon, B. C., & Kruuk, L. E. B. (2010). The misuse of BLUP in ecology and evolution. *American Naturalist*, 175, 116–125.
- Holtmann, B., Lagisz, M., & Nakagawa, S. (2017). Metabolic rates, and not hormone levels, are a likely mediator of between-individual differences in behaviour: A meta-analysis. *Functional Ecology*, 31, 685–696.
- Houslay, T. M., & Wilson, A. J. (2017). Avoiding the misuse in BLUP in behavioral ecology. *Behavioral Ecology*, 28, 948–952.
- Martin, J. S., Massen, J. J. M., Slipogor, V., Bugnyar, T., Jaeggi, A. V., & Koski, S. E. (2019). The EGA plus GNM framework: An integrative approach to modelling behavioural syndromes. *Methods in Ecology and Evolution*, 10, 245–257.
- Martin, J. G. A., & Réale, D. (2008). Temperament, risk assessment and habituation to novelty in eastern chipmunks, *Tamias striatus*. *Animal Behaviour*, 75, 309–318.
- Mitchell, D., Dujon, A. M., Beckmann, C., & Biro, P. A. (2019). Temporal autocorrelation: A neglected factor in the study of behavioral repeatability and plasticity. *Behavioral Ecology*, 31, 222–231.
- Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews*, 85, 935–956.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142.
- Niemelä, P. T., & Dingemanse, N. J. (2017). Individual versus pseudo-repeatability in behaviour: Lessons from translocation experiments in a wild insect. *Journal of Animal Ecology*, 86, 1033–1043.
- Niemelä, P. T., & Dingemanse, N. J. (2018a). Meta-analysis reveals weak associations between intrinsic state and personality. *Proceedings of the Royal Society of London Series B*, 1873, 20172823.
- Niemelä, P. T., & Dingemanse, N. J. (2018b). On the usage of single measurements in behavioural ecology research on individual differences. *Animal Behavior*, 145, 99–105.
- Rausher, M. D. (1992). The measurement of selection on quantitative traits - biases due to environmental covariances between traits and fitness. *Evolution*, 46, 616–626.
- Réale, D., Reader, S. M., Sol, D., McDougall, P., & Dingemanse, N. J. (2007). Integrating temperament in ecology and evolutionary biology. *Biological Reviews*, 82, 291–318.
- Royauté, R., Hedrick, A., & Dochtermann, N. A. (2020). Behavioural syndromes shape evolutionary trajectories via conserved genetic architecture. *Proceedings of the Royal Society B-Biological Sciences*, 287, 20200183.
- Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allee, H., Teplitsky, C., ... Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*. <https://doi.org/10.1111/2041-210X.13434>. In Press.
- Stamps, J. A., & Frankenhuis, W. E. (2016). Bayesian models of development. *Trends in Ecology & Evolution*, 31, 260–268.
- Stinchcombe, J. R., Rutter, M. T., Burdick, D. S., Tiffin, P., Rausher, M. D., & Mauricio, R. (2002). Testing for environmentally induced bias in phenotypic estimates of natural selection: Theory and practice. *American Naturalist*, 160, 511–523.
- Westneat, D. F., Hatch, M. I., Wetzell, D. P., & Ensminger, A. L. (2011). Individual variation in parental care reaction norms: Integration of personality and plasticity. *American Naturalist*, 178, 652–667.
- Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., ... Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology*, 79, 13–26.

**How to cite this article:** Dingemanse NJ, Wright J. Criteria for acceptable studies of animal personality and behavioural syndromes. *Ethology*. 2020;126:865–869. <https://doi.org/10.1111/eth.13082>