Repeatability for time-to-event data

Consistent individual differences in behavior among individuals can have significant ecological and evolutionary consequences (Wolf and Weissing 2012; Dochtermann and Dingemanse 2013). For example, variation in dispersal, range expansion and invasion success can impact gene flow, adaptation and conservation (Cote et al. 2010) with evidence that cognitively flexible (Wright et al. 2010), bold or aggressive (Duckworth and Badyaev 2007) individuals are most likely to push into new environments. Specifically, individual trait variation can impact ecology and evolution when behavioral traits are relatively fixed within individuals either genetically or during development such that individual behavior is consistent across time and distinct from the behavior of other individuals (Reale et al. 2007). To address questions about consistent individual differences, we cannot use identical methodological and statistical approaches as for evaluating population and species average trait values because estimating consistency in individual behavior requires repeated measures within and across individuals (Dingemanse and Dochtermann 2013). Consequently, researchers have developed methods and statistical approaches to meet the growing interest in quantifying consistent individual differences in behavior (Dingemanse and Wright 2020). Despite significant progress, we still do not have a statistically rigorous method for quantifying individual differences from one of the most common forms of behavior data. Time-to-event data, such as latency to approach a novel object, include right-censored values when an individual fails to accomplish the event that make traditional modeling frameworks inappropriate. My aim in applying for this grant is to form a collaboration to development quantitative methods to fill this gap.

Currently, the predominant approach to statistically identify and compare consistent individual differences is to use variance partitioning through mixed-effect models (Dingemanse and Dochtermann 2013). This approach involves evaluating variation in performance among individuals in the population (individual random effect variance) relative to the consistency of individual performance across time or contexts (amount of residual variance). The ratio of among individual variance to total variance results in a “repeatability” estimate (Nakagawa and Schielzeth 2010). Researchers can then determine whether the estimate is significantly different from 0, to indicate there are non-random among-individual differences in behavior, through data permutations where measured values are randomized across individuals to compare the observed repeatability value with a distribution of similar values from randomized data (Nakagawa and Schielzeth 2010). Covariates can also be used in these models to adjust for potentially confounding variables such as body size, sex, or time. Using this method, researchers have identified repeatability in a wide range of behaviors representing personality traits like boldness (e.g., closest approach distance to a novel object: McCune et al. 2018) or exploration (e.g., number of visits to new substrates: Kozlovsky et al. 2014) and cognitive traits like innovation (e.g., ability to extract food from a novel apparatus: Cole et al. 2011). As quantifying repeatability requires estimates of both the residual and among-individual variance components, researchers are currently limited to modeling approaches that allow estimation of both variance components (Oliveira et al. 2016). If the response variable a researcher collects is not well-represented by these distributions, they may be forced to quantify repeatability from a poorly fitting model.

One of the most common types of data collected by animal behavior researchers measures the time until an individual accomplishes something. These types of measures are broadly applicable across many domains of animal behavior. For example, it is possible to compare the latency to disperse from the natal habitat, find a new food source, mount an aggressive response to a simulated intruder, or to initiate foraging or reproductive behavior. In animal cognition studies, researchers frequently measure the time, or the number of trials, until an individual reaches a pre-set criterion or solves a problem (e.g., Griffin and Diquelou 2015). Animal personality researchers often compare the latency for individuals to approach a novel or threatening object or to enter a specific testing arena (e.g., Carter et al. 2012; Perals et al. 2017). To examine repeatability from such time-to-event response variables, researchers may need to assign the maximum trial time value (ceiling value) to individuals that never accomplished the criterion, then analyze these data using random effects models that assume a Poisson or Gaussian distribution (e.g., Lukas et al. 2021; Peignier et al. 2022; Johnstone and Garvey 2023). However, even if the response variables can technically be made to fit within this modeling framework, it is potentially problematic to assign incorrect response values to the individuals that never met the criterion. Logistically, it is not feasible to give all individuals unlimited trials or time during a trial, but individuals that did not meet the criterion within the experimental timeline should be considered as “right-censored” data (Machin et al. 2006). In other words, it was unknown when or if that individual would have accomplished the target behavior. Failing to account for the right-censored nature of these data could bias results as the full range of performance is not known (e.g., van den Heuvel et al. 2023). As part of the collaboration proposed here, we would test the hypothesis that Gaussian and Poisson repeatability models perform differently on complete and censored data.

Survival analysis has largely been used in the medical and insurance fields to evaluate efficacy of medical interventions or the impact of an illness on time until death (Machin et al. 2006). As such, these models specifically account for right-censored data because a reasonable timeframe for most studies does not allow all subjects to be followed until death. Researchers in animal behavior are beginning to use survival analysis for modeling factors related to certain time-to-event behaviors. For example, survival analysis has been used to analyze the influence of social information on the time until naïve individuals solved a novel task (Logan et al. 2016; McCune et al. 2022), the difference in the time it took to solve a foraging task between captive and wild individuals (McCune et al. 2019), and the latency to approach novel objects in different seasons (Greggor et al. 2016). Random effects can be incorporated into survival models to extract estimates of among-individual variance, but it is not currently possible to extract residual variance (Therneau 2022). Thus, despite the frequency with which time-to-event variables are used to measure consistent individual differences we do not yet have a way to use survival models for assessing repeatability. Critically, this leaves the researcher with a choice between two imperfect approaches: utilize a survival model that appropriately accounts for the right-censored nature of the data but does not allow for quantification of repeatability or assess repeatability using a supported modelling approach but with data that include ceiling values and ignore the right-censored nature of the data.

I propose to collaborate with SQuID member Dr. Shinichi Nakagawa to develop a statistical tool for quantifying repeatability from time-to-event response variables. Dr. Nakagawa was integral in developing the rptR package (Stoffel et al. 2017) in the statistical computing program R (R Core Team 2023), which is a user-friendly tool for researchers to quantify repeatability from data modeled with Gaussian, Poisson, binomial, or logistic regression. The objectives of our collaboration will be to 1) determine methods for variance partitioning from mixed-effects survival models or similar analytical techniques that account for censored data; 2) develop the capacity of the rptR package for integrating survival analysis. For this piece, we will also involve SQuID member Dr. Holger Schielzeth, as well as Dr. Martin Stoffel; and 3) the publication of the methods for quantifying individual differences from time-to-event data and a comparison of the performance of this new method relative to results from the current approach of fitting data with ceiling values into a Poisson or Gaussian model. An initial step could be to test the validity of methods that use generalized linear regressions to quantify repeatability, but the dependent variables are pseudo-observations from a Kaplan-Meier survival curve to account for right-censored data (Andersen and Pohar Perme 2010; van den Heuvel et al. 2023). Secondly, we may be able to conduct a survival analysis within a Bayesian framework using a Weibull distribution to estimate residual and among-individual variance from the distribution of the posteriors. We expect the outcome of this collaboration will be a methodological tool that is greatly needed in the toolkit for statistical quantification of individual differences and will be widely applicable in the field of animal behavior.

**Literature cited:**

Andersen, P. K., and M. Pohar Perme. 2010. Pseudo-observations in survival analysis. Statistical Methods in Medical Research 19:71–99.

Carter, A. J., H. H. Marshall, R. Heinsohn, and G. Cowlishaw. 2012. How not to measure boldness: Novel object and antipredator responses are not the same in wild baboons. Animal Behaviour 84:603–609.

Cole, E. F., D. L. Cram, and J. L. Quinn. 2011. Individual variation in spontaneous problem-solving performance among wild great tits. Animal Behaviour 81:491–498.

Cote, J., J. Clobert, T. Brodin, S. Fogarty, and A. Sih. 2010. Personality-dependent dispersal: Characterization, ontogeny and consequences for spatially structured populations. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 365:4065–4076.

Dingemanse, N. J., and N. A. Dochtermann. 2013. Quantifying individual variation in behaviour: Mixed-effect modelling approaches. Journal of Animal Ecology 82:39–54.

Dingemanse, N. J., and J. Wright. 2020. Criteria for acceptable studies of animal personality and behavioural syndromes. Ethology 126:865–869.

Dochtermann, N. A., and N. J. Dingemanse. 2013. Behavioral syndromes as evolutionary constraints. Behavioral Ecology 24:806–811.

Duckworth, R. A., and A. V. Badyaev. 2007. Coupling of dispersal and aggression facilitates the rapid range expansion of a passerine bird. Proceedings of the National Academy of Sciences 104:1–6.

Greggor, A. L., J. W. Jolles, A. Thornton, and N. S. Clayton. 2016. Seasonal changes in neophobia and its consistency in rooks: the effect of novelty type and dominance position. Animal Behaviour 121:11–20.

Griffin, A. S., and M. C. Diquelou. 2015. Innovative problem solving in birds: A cross-species comparison of two highly successful passerines. Animal Behaviour 100:84–94.

Johnstone, K. C., and P. M. Garvey. 2023. To boldly go: methods to quantify personality in mustelids. Animal Behaviour 202:139–147.

Kozlovsky, D. Y., C. L. Branch, C. A. Freas, and V. V. Pravosudov. 2014. Elevation-related differences in novel environment exploration and social dominance in food-caching mountain chickadees. Behavioral Ecology and Sociobiology 68:1871–1881.

Logan, C. J., A. J. Breen, A. H. Taylor, R. D. Gray, and W. Hoppitt. 2016. How New Caledonian crows solve novel foraging problems and what it means for cumulative culture. Learning & Behavior 44:18–28.

Lukas, J., G. Kalinkat, F. W. Miesen, T. Landgraf, J. Krause, and D. Bierbach. 2021. Consistent Behavioral Syndrome Across Seasons in an Invasive Freshwater Fish. Frontiers in Ecology and Evolution 8.

Machin, D., Y. B. Cheung, and M. Parmar. 2006. Survival analysis: a practical approach. John Wiley & Sons.

McCune, K. B., P. G. Jablonski, S. Lee, and R. R. Ha. 2018. Evidence for personality conformity, not social niche specialization in social jays. Behavioral Ecology 29:910–917.

———. 2019. Captive jays exhibit reduced problem-solving performance compared to wild conspecifics. Royal Society Open Science 6:181311.

McCune, K. B., J. J. Valente, P. G. Jablonski, S. Lee, and R. R. Ha. 2022. Social behavior mediates the use of social and personal information in wild jays. Scientific Reports 12:2494.

Nakagawa, S., and H. Schielzeth. 2010. Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. Biological Reviews 85:935–956.

Oliveira, I. R. C., G. Molenberghs, C. G. B. Demétrio, C. T. S. Dias, S. R. Giolo, and M. C. Andrade. 2016. Quantifying intraclass correlations for count and time-to-event data. Biometrical journal. Biometrische Zeitschrift 58:852–867.

Peignier, M., Y. G. Araya-Ajoy, L. Bégué, S. Chaloupka, K. Dellefont, C. Leeb, P. Walsh, et al. 2022. Exploring links between personality traits and their social and non-social environments in wild poison frogs. Behavioral Ecology and Sociobiology 76.

Perals, D., A. S. Griffin, I. Bartomeus, and D. Sol. 2017. Revisiting the open-field test: what does it really tell us about animal personality? Animal Behaviour 123:69–79.

R Core Team. 2023. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Reale, D., S. M. Reader, D. Sol, P. T. McDougall, and N. J. Dingemanse. 2007. Integrating animal temperament within ecology and evolution. Biological Reviews 82:291–318.

Stoffel, M. A., S. Nakagawa, and H. Schielzeth. 2017. rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods in Ecology and Evolution 8:1639–1644.

Therneau, T. M. 2022. Coxme: Mixed Effects Cox models. CRAN.

van den Heuvel, K., J. L. Quinn, A. Kotrschal, and K. van Oers. 2023. Artificial selection for reversal learning reveals limited repeatability and no heritability of cognitive flexibility in great tits (Parus major). Proceedings of the Royal Society B: Biological Sciences 290.

Wolf, M., and F. J. Weissing. 2012. Animal personalities: Consequences for ecology and evolution. Trends in Ecology and Evolution 27:452–461.

Wright, T. F., J. R. Eberhard, E. A. Hobson, M. L. Avery, and M. A. Russello. 2010. Behavioral flexibility and species invasions: The adaptive flexibility hypothesis. Ethology Ecology and Evolution 22:393–404.