Title: Avoiding the misuse of BLUP in behavioral ecology

Thomas M. Houslay*[1], Alastair J. Wilson[1]

[1] Centre of Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Penryn Campus, Penryn, Conrwall, TR10 9FE, United Kingdom.

*Corresponding author:

Email: t.houslay@exeter.ac.uk

Tel: +44(0)7886865020

Running title: Avoiding misuse of BLUP

FUNDING

This work was supported by a Biotechnology and Biological Sciences Research Council grant (BB/L022656/1).

16 ACKNOWLEDGEMENTS

17 We are grateful to members of the Behaviour and Evolution research groups at
18 the University of Exeter's Penryn campus for discussion and for comments on
19 earlier drafts (in particular, T. Bodey, H. Marshall, N. Royle, A. Thornton, T.
20 Tregenza, S. White, and A. Young). Special thanks to E. Wood for illustrating our
21 'study species' in the tutorials. The authors were supported by a Biotechnology
22 and Biological Sciences Research Council grant (BB/L022656/1) awarded to
23 AJW.

24

Lay summary: Research of causes and consequences of animal personality promises exciting insights, yet widely-used tests can lead to spurious results: when predictions of individual-level random effects are used in secondary analyses, their error is not carried forward, leading to increased likelihood of 'false positive' errors. We demonstrate how alternative approaches enable behavioural ecologists to test hypotheses about the causes and consequences of individual behavioural variation while accounting for the uncertainty inherent in the random effects.

Title: Avoiding the misuse of BLUP in behavioral ecology

Running title: Avoiding misuse of BLUP

Abstract: Having recognized that variation around the population-level 'Golden Mean' of labile traits contains biologically meaningful information, behavioral ecologists have focused increasingly on exploring the causes and consequences of individual variation in behavior. These are exciting new directions for the field, assisted in no small part by the adoption of mixed-effects modeling techniques that enable the partitioning of among- and within-individual behavioral variation. It has become commonplace to extract predictions of individual random effects from such models for use in subsequent analyses (for example, between a personality trait and other individual traits such as cognition, physiology, or fitness-related measures). However, these predictions are made with large amounts of error that is not carried forward, rendering further tests susceptible to spurious P-values from these individual-level point

50 estimates. We briefly summarize the problems with such statistical methods that
51 are used regularly by behavioral ecologists, and highlight the robust solutions
52 that exist within the mixed model framework, providing tutorials to aid in their
53 implementation.
54
55

Characterizing individual variation in behavior is an exciting research area in behavioral ecology, with great interest in the fields of 'animal personality' and individual differences in behavioral plasticity (Réale, Dingemanse, et al. 2010; Japyassú and Malange 2014). This research is predicated on exploring previously ignored phenotypic variation: behavioral ecologists have escaped the 'tyranny of the Golden Mean' in labile traits (Bennett 1987; Wilson 1998; Williams 2008), and are increasingly finding meaningful biology in what was formerly considered residual variation (Cleasby and Nakagawa 2011; Stamps et al. 2012; Brommer 2013a). Progress in these fields has been boosted by the adoption of mixed-effects modeling techniques, particularly the use of quantitative genetics-style approaches for partitioning phenotypic variation into its 'between-individual' and 'within-individual' components (Nussey et al. 2007; Smiseth et al. 2008; van de Pol and Wright 2009; Dingemanse et al. 2012; Dingemanse and Dochtermann 2013; Royle et al. 2014; Allegue et al. 2016). Behavioral ecologists are also increasingly interested in extending these analyses of individual behavioral variation for new avenues and purposes (Sih et al. 2004; Dall et al. 2012; Japyassú and Malange 2014; Roche et al. 2016; Stamps and Biro 2016). These typically involve exploration of the causes and consequences of individual variation in behavior (and/or behavioral plasticity), by testing for their association with variation in other individual traits (e.g., physiological, cognitive, social, or fitness-related) or environmental variables. However, the use of anticonservative methods has become pervasive in this field. Here we highlight not only the problems with a widely-used approach in the study of individual behavioral variation, but also the straightforward statistical solutions to these problems that should thereby hasten progress.

Specifically, it has become common practice to extract predictions of individual random effects from fitted mixed models and to use these in subsequent analyses, such as correlation tests or linear regression models (Table 1). Problems arise from this approach because individual point estimates from random effects in mixed models (sometimes known as conditional modes, or best linear unbiased predictors, BLUPs) are predicted with large amounts of error. Their use in secondary analyses can therefore lead to highly anticonservative tests of biological hypotheses, because the error inherent in their prediction is excluded from these further tests (Hadfield et al. 2010). We stress that BLUP is an incredibly useful technique that should not be dismissed in any way as inherently 'bad' (Robinson 1991). Indeed, it is entirely appropriate to use individual-level predictions to say something about individuals (or genotypes, or specific levels of some other random effect). For example, scrutiny of BLUPs could be used to identify which individuals are the 'boldest', or to select individuals for groups to be used in further experimental study. However, when the objective is to say something about population-level processes or relationships then analyzing sets of model predictions while ignoring their associated error is not statistically correct. This has been recognized in other fields (notably ecological and evolutionary quantitative genetics), but less so in behavioral ecology, where these improper analyses persist. As detailed by Hadfield et al. (2010), such analyses can therefore result in spuriously narrow confidence intervals and/or spuriously low P-values that are interpreted as indicators of biological significance. While the qualitative conclusions of individual papers employing these methods may prove robust in many cases,

failure to properly account for uncertainty will increase Type I errors (false positives) across the field. In short, published P-values are systematically anticonservative and should not be taken at face value.

Recent examples of publications (mis-)using BLUPs include tests of associations between personality (and/or individual variation in behavioral plasticity) and a wide range of traits, including physiology, cognition, social networks, niche specialization, and fitness (Table 1). In many cases, authors have explicitly acknowledged the potential for problems as outlined by Hadfield *et al*. (2010). Nonetheless, use of these 'stats on stats' approaches that are known to be inappropriate for hypothesis testing (see Brommer 2013b for further discussion) continues unabated. This is presumably because researchers are not aware of how to implement more robust analytical strategies, and/or because of a misconception that problems are restricted to quantitative genetic models. On the latter point we note that predictions from mixed models in which random effects are assumed to covary between individuals (through e.g., genetic relatedness, spatial/temporal autocorrelation, or social processes) cannot be treated as independent 'data points'. However, this in no way justifies ignoring uncertainty when random effects are predicted from a model that assumes no among-individual covariance.

Fortunately, the mixed-effects model framework does offer a way to test hypotheses such as those listed above while fully accounting for the uncertainty inherent in the random effects. An overreliance on the (otherwise excellent) lme4 package for mixed models in R (Bates et al. 2015) may have held many

131 behavioral ecologists in the 'Flatland' of univariate modeling (Walsh 2007). In
132 the majority of cases, questions that are multivariate in nature are best answered
133 using a multivariate framework. That is, a modeling framework containing
134 multiple response variables, enabling (i) testing of how explanatory variables
135 ('fixed effects') predict these responses, as in standard univariate models, and
136 (ii) the simultaneous estimation of the variance of each response and the
137 covariance between them, at group levels specified within the random effects
138 structure. It is relatively straightforward to rephrase these multivariate
139 questions in terms of variances and covariances (or derived correlations and
140 regressions), and to fit multivariate models accordingly (some examples include
141 Ferrari et al. 2013; Kluen et al. 2013; Royauté et al. 2013; Boulton et al. 2014;
142 Careau et al. 2015; Niemelä et al. 2015; Petelle et al. 2015; Sanderson et al. 2015;
143 Santostefano et al. 2016; Vallon et al. 2016; White et al. 2016). For instance, we
144 might hypothesize a behavioral syndrome in which positive correlations are
145 predicted between the (repeatable) tendencies of individuals to exhibit three
146 behaviors. Having assayed each of these behaviors on multiple occasions for a
147 set of individuals, the correct approach would be to estimate – and test the
148 significance of – those among-individual correlations directly in a trivariate
149 mixed model incorporating all of the behavioral data. This method yields
150 correlation estimates with valid measures of uncertainty (SE or CI). This is not
151 the case when generating individual-level random effect predictions from three
152 separate univariate models (one for each behavior) and then testing whether
153 they are correlated. In the latter approach, uncertainty will be underestimated
154 and thus Type I error is more likely to occur (Figure 1).
155

156 On a pragmatic point, we note that it is not required that each variable of interest
157 be a repeated measure in these models – for example, it is perfectly feasible to
158 test for the existence of an among-individual correlation between a personality
159 trait (with repeated measures) and some other variable with only one
160 observation per individual, such as an estimate of its lifetime reproductive
161 success. In the supplementary material, we provide worked examples of how to
162 set up multivariate statistical models to address these (and several similar)
163 questions using the R packages ASReml-R (Butler 2009) and MCMCglmm
164 (Hadfield 2010). These examples provide users with the tools to test their
165 hypotheses in a multivariate framework, incorporating all of their data and
166 avoiding potentially spurious results.
167
168 We also note that multivariate mixed models may often provide a more
169 appropriate route to testing hypotheses about multivariate phenotypes in other
170 contexts. For instance, one approach to exploring behavioral syndromes has
171 been to reduce the dimensionality of observed behaviors by performing
172 principal components analysis (PCA) on multivariate data, and then to use
173 univariate mixed models to calculate repeatability on individual scores for each
174 component (e.g., Edenbrow & Croft 2013; Le Galliard *et al.* 2013; Brent *et al.*
175 2014; Patrick & Weimerskirch 2014; Sussman *et al.* 2014; Rangassamy *et al.*
176 2015). This allows us to ask whether, for instance, the major axis of observed
177 behavioral (co)variation is repeatable. This is a valid question but in many cases
178 perhaps not the most pertinent one, since the first principal component of
179 observed variation includes both among- and within-individual trait variation.
180 For studies of individual differences in behavior, the more relevant question

might be better focused at the among-individual level – that is, what does the major axis of among-individual variation look like? If so, then isolating the among-individual (co)variance matrix (sometimes denoted **I**; Wilson et al. 2011) by applying a multivariate mixed model to a set of traits is the proper first step. Principal components (or eigen vectors) of **I** can then be examined directly. This strategy is probably more appropriate for testing models such as 'pace of life syndrome' or stress coping styles that posit trait correlations at the among-individual level – i.e., that these correlations are due to consistent differences among individuals, and not because of some temporary aspect of environmental variation (Koolhaas et al. 2007; Carere et al. 2010; Coppens et al. 2010; Réale, Garant, et al. 2010; Carter et al. 2013). The value of partitioning individual (co)variances has been discussed in more detail by Brommer (2013a), and illustrations exist in the literature of the use of multivariate mixed models for studying pace of life syndrome (White et al. 2016) and stress coping styles (Boulton et al. 2015).

We fully acknowledge that multivariate mixed models are data hungry. However, a failure of these multivariate models to converge to sensible and/or precise solutions does not mean that we can retreat to the relative comfort of previous methods: in fact, this is likely to indicate a lack of power to answer the question at hand (see Martin *et al.* 2011; Wolak, Fairbairn & Paulsen 2012). In cases where logistical constraints prevent there being enough measurements to partition out the among-individual behavioral (co)variation, a preferable method may sometimes be to work with observed phenotypic (co)variance while acknowledging this and the assumptions that underpin conclusions drawn.

206 Indeed, much of behavioral ecology is predicated on the 'phenotypic gambit', the
207 assumption that phenotypic patterns of trait (co)variation (denoted **P**) provide a
208 workable proxy for patterns of genetic (co)variance (**G**). If **P** can be used (with
209 caveats) in place of **G** where estimation of genetic parameters is not feasible,
210 then it can also be used (with caveats) in place of **I** where partitioning of among-
211 from within-individual covariation is not feasible.
212
213 To conclude, we absolutely wish to encourage more studies that further our
214 understanding of the causes and consequences of individual differences in
215 behavior. However, we also make a plea to the community to avoid
216 inappropriate methods of analysis that lead to spurious precision and increased
217 Type I errors. This field depends upon embracing the power of previously
218 ignored phenotypic variation, and it is flourishing because of the exciting
219 questions we can now address – but we must ensure that we use the right tools
220 when doing so.
221
222 REFERENCES
223
224 Adriaenssens B, Johnsson JI. 2011. Shy trout grow faster: Exploring links
225 between personality and fitness-related traits in the wild. Behav. Ecol. 22:135–
226 143.
227 Adriaenssens B, Pauliny A, Blomqvist D, Johnsson JI. 2016. Telomere length
228 covaries with personality in wild brown trout. Physiol. Behav. 165:217–222.
229 Allegue H, Araya-Ajoy YG, Dingemanse NJ, Dochtermann NA, Garamszegi LZ,
230 Nakagawa S, Réale D, Schielzeth H, Westneat DF. 2016. SQuID - Statistical

231 Quantification of Individual Differences: an educational and statistical tool for
232 understanding multi-level phenotypic data in linear mixed models. Methods Ecol.
233 Evol.
234 Bates DM, Mächler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects
235 Models Using lme4. J. Stat. Softw. 67.
236 Beckmann C, Biro PA. 2013. On the Validity of a Single (Boldness) Assay in
237 Personality Research. Ethology 119:937–947.
238 Bennett AF. 1987. Interindividual variability: an underutilized resource. New Dir.
239 Ecol. Physiol.:147–169.
240 Bergeron P, Montiglio PO, Réale D, Humphries MM, Gimenez O, Garant D. 2013.
241 Disruptive viability selection on adult exploratory behaviour in eastern
242 chipmunks. J. Evol. Biol. 26:766–774.
243 Betini GS, Norris DR. 2012. The relationship between personality and plasticity
244 in tree swallow aggression and the consequences for reproductive success. Anim.
245 Behav. 83:137–143.
246 Boulton K, Couto E, Grimmer AJ, Earley RL, Canario AVM, Wilson AJ, Walling CA.
247 2015. How integrated are behavioral and endocrine stress response traits? A
248 repeated measures approach to testing the stress-coping style model. Ecol. Evol.
249 5:618–633.
250 Boulton K, Grimmer AJ, Rosenthal GG, Walling CA, Wilson AJ. 2014. How stable
251 are personalities? A multivariate view of behavioural variation over long and
252 short timescales in the sheepshead swordtail, Xiphophorus birchmanni. Behav.
253 Ecol. Sociobiol. 68:791–803.
254 Brent LJN, Semple S, MacLarnon A, Ruiz-Lambides A, Gonzalez-Martinez J, Platt
255 ML. 2014. Personality Traits in Rhesus Macaques (Macaca mulatta) Are Heritable

but Do Not Predict Reproductive Output. Int. J. Primatol. 35:188–209.

Brommer JE. 2013a. On between-individual and residual (co)variances in the study of animal personality: Are you willing to take the “individual gambit”? Behav. Ecol. Sociobiol. 67:1027–1032.

Brommer JE. 2013b. Phenotypic plasticity of labile traits in the wild. Curr. Zool. 59:485–505.

Brust V, Guenther A. 2015. Domestication effects on behavioural traits and learning performance: comparing wild cavies to guinea pigs. Anim. Cogn. 18:99–109.

Butler D. 2009. asreml: asreml() fits the linear mixed model.

Careau V, Montiglio PO, Garant D, Pelletier F, Speakman JR, Humphries MM, Réale D. 2015. Energy expenditure and personality in wild chipmunks. Behav. Ecol. Sociobiol. 69:653–661.

Carere C, Caramaschi D, Fawcett TW. 2010. Covariation between personalities and individual differences in coping with stress: Converging evidence and hypotheses. Curr. Zool. 56:728–741.

Carter AJ, English S, Clutton-Brock TH. 2014. Cooperative personalities and social niche specialization in female meerkats. J. Evol. Biol. 27:815–825.

Carter AJ, Feeney WE, Marshall HH, Cowlishaw G, Heinsohn R. 2013. Animal personality: What are behavioural ecologists measuring? Biol. Rev. 88:465–475.

Carter AJ, Goldizen A, Heinsohn R. 2012. Personality and plasticity: Temporal behavioural reaction norms in a lizard, the Namibian rock agama. Anim. Behav. 84:471–477.

Carter AJ, Heinsohn R, Goldizen AW, Biro PA. 2012. Boldness, trappability and sampling bias in wild lizards. Anim. Behav. 83:1051–1058.

281 Carter AJ, Marshall HH, Heinsohn R, Cowlishaw G. 2012. How not to measure
282 boldness: Novel object and antipredator responses are not the same in wild
283 baboons. Anim. Behav. 84:603–609.
284 Cleasby IR, Nakagawa S. 2011. Neglected biological patterns in the residuals.
285 Behav. Ecol. Sociobiol. 65:2361–2372.
286 Coppens CM, de Boer SF, Koolhaas JM. 2010. Coping styles and behavioural
287 flexibility: towards underlying mechanisms. Philos. Trans. R. Soc. Lond. B. Biol.
288 Sci. 365:4021–4028.
289 Dall SRX, Bell AM, Bolnick DI, Ratnieks FLW, Sih A. 2012. An evolutionary ecology
290 of individual differences. Ecol. Lett. 15:1189–98.
291 Dammhahn M, Almeling L. 2012. Is risk taking during foraging a personality
292 trait? A field test for cross-context consistency in boldness. Anim. Behav. 84:131–
293 1139.
294 Dingemanse NJ, Dochtermann NA. 2013. Quantifying individual variation in
295 behaviour: mixed-effect modelling approaches. van de Pol M, editor. J. Anim.
296 Ecol. 82:39–54.
297 Dingemanse NJ, Dochtermann NA, Nakagawa S. 2012. Defining behavioural
298 syndromes and the role of “syndrome deviation” in understanding their
299 evolution. Behav. Ecol. Sociobiol. 66:1543–1548.
300 Edenbrow M, Croft DP. 2013. Environmental and genetic effects shape the
301 development of personality traits in the mangrove killifish Kryptolebias
302 marmoratus. Oikos 122:667–681.
303 Ferrari C, Pasquaretta C, Carere C, Cavallone E, von Hardenberg A, Réale D. 2013.
304 Testing for the presence of coping styles in a wild mammal. Anim. Behav.
305 85:1385–1396.

306 Finkemeier MA, Trillmich F, Guenther A. 2016. Match-Mismatch Experiments
307 Using Photoperiod Expose Developmental Plasticity of Personality Traits.
308 Ethology 122:80–93.
309 Fuong H, Maldonado-Chaparro A, Blumstein DT. 2015. Are social attributes
310 associated with alarm calling propensity? Behav. Ecol. 26:587–592.
311 Le Galliard JF, Paquet M, Cisel M, Montes-Poloni L. 2013. Personality and the
312 pace-of-life syndrome: Variation and selection on exploration, metabolism and
313 locomotor performances. Funct. Ecol. 27:136–144.
314 Geffroy B, Bru N, Dossou-Gbété S, Tentelier C, Bardonnet A. 2014. The link
315 between social network density and rank-order consistency of aggressiveness in
316 juvenile eels. Behav. Ecol. Sociobiol. 68:1073–1083.
317 Gibelli J, Dubois F. 2016. Does personality affect the ability of individuals to track
318 and respond to changing conditions? Behav. Ecol. 0:arw137.
319 Guenther A, Brust V, Dersen M, Trillmich F. 2014. Learning and personality types
320 are related in cavies (Cavia aperea). J. Comp. Psychol. 128:74–81.
321 Guenther A, Trillmich F. 2015. Within-litter differences in personality and
322 physiology relate to size differences among siblings in cavies. Physiol. Behav.
323 145:22–28.
324 Hadfield JD. 2010. MCMC Methods for Multi-Response Generalized Mixed
325 Models: The MCMCglmm R Package. J. Stat. Softw. 33:1–25.
326 Hadfield JD, Wilson AJ, Garant D, Sheldon BC, Kruuk LEB. 2010. The misuse of
327 BLUP in ecology and evolution. Am. Nat. 175:116–25.
328 Han CS, Brooks RC. 2014. Long-Term Effect of Social Interactions on Behavioral
329 Plasticity and Lifetime Mating Success. Am. Nat. 183:431–444.
330 Hellström G, Heynen M, Borcherding J, Magnhagen C. 2016. Individual

consistency and context dependence in group-size preference of Eurasian perch. Behav. Processes 133:6–11.

Heynen M, Borcherding J, Bunnefeld N, Magnhagen C. 2016. Plasticity and consistency of behavioural responses to predation risk in laboratory environments. J. Zool. 300:228–235.

Japyassú HF, Malange J. 2014. Plasticity, stereotypy, intra-individual variability and personality: Handle with care. Behav. Processes 109:40–47.

Kelley AD, Humphries MM, McAdam AG, Boutin S. 2015. Changes in wild red squirrel personality across ontogeny: activity and aggression regress towards the mean. Behaviour 152:1291–1306.

Kluen E, Siitari H, Brommer JE. 2013. Testing for between individual correlations of personality and physiological traits in a wild bird. Behav. Ecol. Sociobiol. 68:205–213.

Koolhaas JM, de Boer SF, Buwalda B, Van Reenen K. 2007. Individual variation in coping with stress: A multidimensional approach of ultimate and proximate mechanisms. Brain. Behav. Evol. 70:218–226.

Magnhagen C, Hellström G, Borcherding J, Heynen M. 2012. Boldness in two perch populations - long-term differences and the effect of predation pressure. J. Anim. Ecol. 81:1311–1318.

Martin JGA, Nussey DH, Wilson AJ, Réale D. 2011. Measuring individual differences in reaction norms in field and experimental studies: a power analysis of random regression models. Methods Ecol. Evol. 2:362–374.

Montiglio P-O, DiRienzo N. 2016. There's no place like home: the contribution of direct and extended phenotypes on the expression of spider aggressiveness. Behav. Ecol. 27:arw094.

356 Montiglio P-O, Wey TW, Chang AT, Fogarty S, Sih A. 2016. Multiple mating
357 reveals complex patterns of assortative mating by personality and body size.
358 Quinn J, editor. J. Anim. Ecol. 85:125–135.
359 Montiglio P-O, Wey TW, Chang AT, Fogarty S, Sih A. 2016 Nov. Correlational
360 selection on personality and social plasticity: morphology and social context
361 determine behavioural effects on mating success. J. Anim. Ecol.
362 Montiglio PO, Garant D, Pelletier F, Réale D. 2012. Personality differences are
363 related to long-term stress reactivity in a population of wild eastern chipmunks,
364 Tamias striatus. Anim. Behav. 84:1071–1079.
365 Niemelä PT, Lattenkamp EZ, Dingemanse NJ. 2015. Personality-related survival
366 and sampling bias in wild cricket nymphs. Behav. Ecol. 26:936–946.
367 Nussey DH, Wilson AJ, Brommer JE. 2007. The evolutionary ecology of individual
368 phenotypic plasticity in wild populations. J. Evol. Biol. 20:831–44.
369 Patrick SC, Weimerskirch H. 2014. Personality, foraging and fitness
370 consequences in a long lived seabird. PLoS One 9.
371 Petelle MB, Martin JG a, Blumstein DT. 2015. Heritability and genetic correlations
372 of personality traits in a wild population of yellow-bellied marmots (Marmota
373 flaviventris). J. Evol. Biol. 28:1840–1848.
374 Pineaux M, Turgeon J. 2017. Behavioural Consistency in Female Resistance to
375 Male Harassment in a Water Strider Species. Hebets E, editor. Ethology 123:83–
376 93.
377 van de Pol M, Wright J. 2009. A simple method for distinguishing within- versus
378 between-subject effects using mixed models. Anim. Behav. 77:753–758.
379 Rangassamy M, Dalmas M, Féron C, Gouat P, Rödel HG. 2015. Similarity of
380 personalities speeds up reproduction in pairs of a monogamous rodent. Anim.

Behav. 103:7–15.

Réale D, Dingemanse NJ, Kazem AJN, Wright J. 2010. Evolutionary and ecological approaches to the study of personality. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365:3937–3946.

Réale D, Garant D, Humphries MM, Bergeron P, Careau V, Montiglio P-O. 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.

Robinson GK. 1991. That BLUP is a good thing: The estimation of random effects. Stat. Sci. 6:15–32.

Roche DG, Careau V, Binning SA. 2016. Demystifying animal "personality" (or not): why individual variation matters to experimental biologists. J. Exp. Biol.

Royauté R, Buddle CM, Vincent C. 2013. Interpopulation Variations in Behavioral Syndromes of a Jumping Spider from Insecticide-Treated and Insecticide-Free Orchards. Foster S, editor. Ethology 119:1–13.

Royle NJ, Russell AF, Wilson AJ. 2014. The evolution of flexible parenting. Science. 345:776–781.

Sanderson JL, Stott I, Young AJ, Vitikainen EIK, Hodge SJ, Cant MA. 2015. The origins of consistent individual differences in cooperation in wild banded mongooses, Mungos mungo. Anim. Behav. 107:193–200.

Santostefano F, Wilson AJ, Araya-Ajoy YG, Dingemanse NJ. 2016. Interacting with the enemy: Indirect effects of personality on conspecific aggression in crickets. Behav. Ecol. 27:1235–1246.

Schell CJ, Young JK, Lonsdorf E V., Mateo JM, Santymire RM. 2016. Olfactory attractants and parity affect prenatal androgens and territoriality of coyote breeding pairs. Physiol. Behav. 165:43–54.

406 Shonfield J, Taylor RW, Boutin S, Humphries MM, McAdam AG. 2012. Territorial
407 defence behaviour in red squirrels is influenced by local density. Behaviour
408 149:369–390.
409 Sih A, Bell AM, Johnson JC. 2004. Behavioral syndromes: an ecological and
410 evolutionary overview. Trends Ecol. Evol. 19:372–8.
411 Smiseth PT, Wright J, Kölliker M. 2008. Parent-offspring conflict and co-
412 adaptation: behavioural ecology meets quantitative genetics. Proc. Biol. Sci.
413 275:1823–30.
414 Stamps JA, Biro PA. 2016. Personality and individual differences in plasticity.
415 Curr. Opin. Behav. Sci. 12:18–23.
416 Stamps JA, Briffa M, Biro PA. 2012. Unpredictable animals: Individual differences
417 in intraindividual variability (IIV). Anim. Behav. 83:1325–1334.
418 Sussman AF, Mates EA, Ha JC, Bentson KL, Crockett CM. 2014. Tenure in current
419 captive setting and age predict personality changes in adult pigtailed macaques.
420 Anim. Behav. 89:23–30.
421 Vallon M, Grom C, Kalb N, Sprenger D, Anthes N, Lindström K, Heubel KU. 2016.
422 You eat what you are: Personality-dependent filial cannibalism in a fish with
423 paternal care. Ecol. Evol. 6:1340–1352.
424 Walsh B. 2007. Escape from flatland. J. Evol. Biol. 20:36–38.
425 Wey TW, Chang AT, Fogarty S, Sih A. 2014. Personalities and presence of
426 hyperaggressive males influence male mating exclusivity and effective mating in
427 stream water striders. Behav. Ecol. Sociobiol. 69:27–37.
428 Wey TW, Chang AT, Montiglio P-O, Fogarty S, Sih A. 2015. Linking short-term
429 behavior and personalities to feeding and mating rates in female water striders.
430 Behav. Ecol. 26:1196–1202.

White SJ, Kells TJ, Wilson AJ. 2016. Metabolism, personality and pace of life in the Trinidadian guppy, Poecilia reticulata. Behaviour 153:1517–1543.

Williams TD. 2008. Individual variation in endocrine systems: moving beyond the “tyranny of the Golden Mean”. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 363:1687–1698.

Wilson AJ, de Boer M, Arnott G, Grimmer A. 2011. Integrating Personality Research and Animal Contest Theory: Aggressiveness in the Green Swordtail Xiphophorus helleri. Briffa M, editor. PLoS One 6:e28024.

Wilson DS. 1998. Adaptive individual differences within single populations. Philos. Trans. R. Soc. B Biol. Sci. 353:199–205.

Wolak ME, Fairbairn DJ, Paulsen YR. 2012. Guidelines for estimating repeatability. Methods Ecol. Evol. 3:129–137.

Wuerz Y, Krüger O. 2015. Personality over ontogeny in zebra finches: long-term repeatable traits but unstable behavioural syndromes. Front. Zool. 12:S9.

447 FIGURE LEGENDS

448

449 Figure 1: Taken from a worked example provided in the Supplementary
450 Information, (a) shows a scatterplot of individual-level estimates (BLUPs) of two
451 personality traits, extracted from separate univariate models. Bars around each
452 point show the standard error of the estimate for both traits, which is ignored by
453 subsequent analyses of these BLUPs. Testing a correlation using only BLUPs and
454 ignoring their error results in an anticonservative test, as illustrated in (b). The
455 correlation test using BLUPs produces narrow confidence intervals, and a
456 correspondingly small P-value of 0.0019, indicating statistical significance
457 ('BLUP' on x-axis). However, testing the correlation directly in a bivariate model
458 using REML and retaining all data returns larger (approximate) confidence
459 intervals which straddle zero (95% CI approximated as r +/- 1.96SE) and a P-
460 value (based on a likelihood ratio test) of 0.12, such that the correlation is not
461 statistically significant ('Bivariate ASReml' on x-axis). Using the same data,
462 Bayesian 95% credible intervals also cross zero, which indicates a lack of
463 statistical significance ('Bivariate MCMCglmm').

464

465 TABLES

466

467 Table 1: Examples in the behavioural literature of questions regarding individual
468 variation in behaviour ('personality') and behavioural plasticity, using best linear
469 unbiased predictors (BLUPs) in secondary analyses rather than multivariate
470 models. All were published after the publication of Hadfield *et al* (2010).

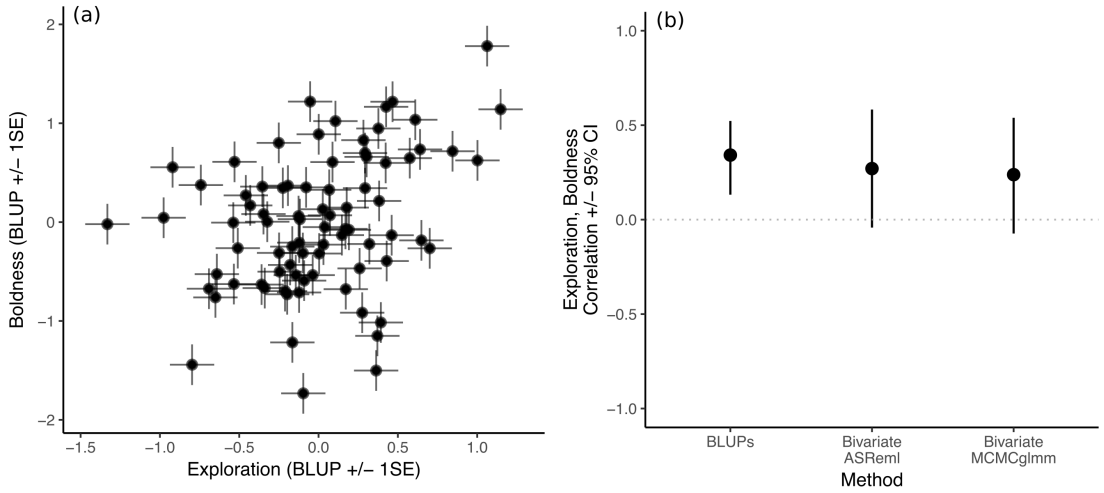
Test	Species	Reference
Behavioural syndromes	*Taeniopygia guttata*	(Wuerz and Krüger 2015)
	Latrodectus hesperus	(Montiglio and DiRienzo 2016)
Personality across life stages	*Tamiasciurus hudsonicus*	(Kelley et al. 2015)
Different measures of a single personality trait	*Bitis arietans*	(Carter, Marshall, et al. 2012)
	Pomacentrus wardi, P. amboinensis	(Beckmann and Biro 2013)
Personality & sampling bias	*Agama planiceps*	(Carter, Heinsohn, et al. 2012)
Personality & hormones	*Tamias striatus*	(Montiglio et al. 2012)
	Canis latrans	(Schell et al. 2016)
Personality & physiology	*Cavia aperea*	(Guenther and Trillmich 2015)
	C. aperea	(Finkemeier et al. 2016)
Personality & telomere length	*Salmo trutta*	(Adriaenssens et al. 2016)
Personality & cognition	*C. aperea*	(Guenther et al. 2014)
	C. aperea, C. porcellus	(Brust and Guenther 2015)
Personality & social network attributes	*Anguilla anguilla*	(Geffroy et al. 2014)
	Marmota flaviventris	(Fuong et al. 2015)
Personality & local density	*T. hudsonicus*	(Shonfield et al. 2012)
Personality & social niche specialisation	*Suricata suricatta*	(Carter et al. 2014)
Personality & group-size preference	*Perca fluviatilis*	(Hellström et al. 2016)
Personality & predation risk	*P. fluviatilis*	(Magnhagen et al. 2012) (Heynen et al. 2016)
Personality & mating behaviour	*Aquarius remigis*	(Wey et al. 2014; Wey et al. 2015)
	Gerris buenoi	(Pineaux and Turgeon 2016)
Personality & survival	*T. striatus*	(Bergeron et al. 2013)
Personality & fitness-related traits	*S. trutta*	(Adriaenssens and Johnsson 2011)
Personality & individual variation in behavioural plasticity	*A. planiceps*	(Carter, Goldizen, et al. 2012)
	Microcebus murinus	(Dammhahn and Almeling 2012)
	T. guttata	(Gibelli and Dubois 2016)
Personality, behavioural plasticity & reproductive success	*Tachycineta bicolor*	(Betini and Norris 2012)
Personality, behavioural plasticity & mating	*A. remigis*	(Montiglio et al. 2016a; Montiglio et al. 2016b)
Personality, behavioural plasticity & fitness	*Tenagogerris euphrosyne*	(Han and Brooks 2014)

471

472

473 FIGURES

474 Figure 1



475

Avoiding the misuse of BLUP in behavioral ecology: Multivariate modelling for individual variation (ASReml-R tutorial)

T.M. Houslay & A.J. Wilson, Behavioral Ecology

January 2017

Introduction

Overview

This tutorial accompanies our paper, "Avoiding the misuse of BLUP in behavioral ecology". Below, we provide worked examples of multivariate statistical methods for directly testing hypotheses about associations between individual variation in behaviour and other traits. Below, we will:

- Test the correlation between two personality traits (behaviours measured repeatedly on individuals);
- Test for an association between these personality traits and a measure of fitness (one value per individual).

In this version, we illustrate these models using the `R` interface for `ASReml`, which is commercial software available from VSNi. We have provided a separate tutorial for the free `R` package `MCMCglmm`, but note that `MCMCglmm` uses Bayesian methods while `ASReml` uses maximum likelihood (and is therefore likely to be more familiar to users of the `R` package `lme4`).

Aims

Please note that we do assume readers are familiar with the general principles of specifying univariate mixed effects models, and using diagnostic plots to check that the fitted model does not violate assumptions of the linear model. Readers unfamiliar with using univariate mixed effects models for modelling a single behavioural trait might prefer to start with (for example) Dingemanse & Dochtermann's 2013 paper, 'Quantifying individual variation in behaviour: mixed effects modelling approaches'.

We also use various methods for manipulating and visualising data frames using the `tidyverse` package (including `tidyr`, `dplyr`, `ggplot2` etc) — more details on their use can be found at http://r4ds.had.co.nz/.

In our tutorial, we aim to teach the following:

- How to phrase questions of interest in terms of variances and covariances (or derived correlations or regressions);
- How to incorporate more advanced model structures, such as:
  - Fixed effects that apply only to a subset of the response traits;
  - Traits which are measured a different number of times (*e.g.*, repeated measures of behaviour and a single value of breeding success);
- Hypothesis testing using likelihood ratio tests.

Packages required

There are several packages that you must have installed in `R` prior to starting this tutorial:

- `asreml` (note that this should be provided by the vendor, VSNi)
- `lme4`
- `nadiv`
- `tidyverse`
- `broom`

'Study system'

For this tutorial, we have collected data on a population of wild haggis (*Haggis scoticus*) that roam the Highlands of Scotland.

Figure 1: A male haggis in the wild (thanks to Emma Wood, http://www.ewood-art.co.uk/)

We tag all haggis individually when they emerge from their burrows as juveniles in their first spring. Here, we concentrate on male haggis, which are solitary and territorial. Previous work has identified behaviours that can be measured repeatedly, and used to represent the personality traits **boldness** and **exploration**. We also have the ability to collect a single measure of mating success (as a fitness proxy) for each male at the end of the season.

Behavioural syndromes

One type of 'behavioural syndrome' is a correlation between personality traits. Since personality can be viewed (under most definitions) as the repeatable (among-individual) component of behaviour, evidence

for the presence of a behavioural syndrome is provided by covariance among behaviours that arises from among-individual differences.

Here we have repeatedly measured behaviours that represent **boldness** and **exploration**. We observed each behaviour 4 times per individual. We also measured their body size on the day of behavioural assay so as to control for general size effects in our statistical models.

Load libraries and inspect data

```
library(lme4)
library(asreml)
library(tidyverse)
library(broom)
library(nadiv)

df_syndrome <- read_csv("syndrome.csv")
```

This data frame has 6 variables:

- Individual **ID**
- The repeat number for each behavioural test, **assay_rep**
- **boldness**, measured 4 times per individual
- **exploration**, measured 4 times per individual
- **fitness**, our measure of mating success, with a single value for each individual
- Individual **body_size**, as measured on the day of testing.

Univariate models

We first use the `R` package `lme4` to determine the proportion of phenotypic variation (adjusted for fixed effects) that is due to differences among individuals, separately for each behaviour. We assume readers have knowledge of these 'univariate' models and their use in behavioural studies — if not, there are various other publications that go into them in greater detail (e.g., Dingemanse & Dochtermann (2013)).

Boldness

Our model includes fixed effects of the assay repeat number (centred) and individual body size (centred and scaled to standard deviation units), as we wish to control for any systematic effects of these variables on individual behaviour. Please be aware that controlling variables are at your discretion — for example, while we want to characterise among-individual variance in boldness after controlling for size effects in this study, others may wish to characterise among-individual variance in boldness without such control. Indeed, using the techniques shown later in this tutorial, it would be entirely possible to characterise both among-individual variance in boldness and in size, and the among-individual covariance between these measurements.

```
lmer_b <- lmer(boldness ~ scale(assay_rep, scale=FALSE) +
                 scale(body_size) +
                 (1|ID),
               data = df_syndrome)
plot(lmer_b)
qqnorm(residuals(lmer_b))
hist(residuals(lmer_b))

summary(lmer_b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: boldness ~ scale(assay_rep, scale = FALSE) + scale(body_size) +
##     (1 | ID)
##    Data: df_syndrome
##
## REML criterion at convergence: 1061.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.3645 -0.6496 -0.1154  0.6463  2.6894
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.6951   0.8337
##  Residual             1.1682   1.0808
## Number of obs: 320, groups:  ID, 80
##
## Fixed effects:
##                                 Estimate Std. Error t value
## (Intercept)                     20.09133    0.11108  180.87
## scale(assay_rep, scale = FALSE) -0.04805    0.05404   -0.89
## scale(body_size)                 0.14128    0.10893    1.30
##
## Correlation of Fixed Effects:
##             (Intr) s(_s=F
## s(_,s=FALSE  0.000
## scl(bdy_sz)  0.000 -0.002
```

Having examined diagnostic plots of the model fit, we can check the model `summary`. We are interested in the *random effects* section of the `lme4` model output (specifically the **variance** component — note that the standard deviation here is simply the square root of the variance). Evidence for 'animal personality' (or 'consistent among-individual differences in behaviour') in the literature is largely taken from the **repeatability** of behaviorual traits: we can compute this **repeatability** (also known as the *intraclass correlation coefficient*) by dividing the variance in the trait due to differences among individuals (V_{ID}) by the total phenotypic variance after accounting for the fixed effects ($V_{ID} + V_{residual}$). This can be done quickly and automatically through the use of the R package `broom`:

```
rep_bold <- tidy(lmer_b, effects = "ran_pars", scales = "vcov") %>%
  select(group, estimate) %>%
  spread(group, estimate) %>%
  mutate(repeatability = ID/(ID + Residual))

rep_bold
```

ID	Residual	repeatability
0.695	1.168	0.373

So we can see that 37.3% of the phenotypic variation in boldness (having controlled for body size and assay repeat number) is due to differences among individuals.

Let's do the same for our other behavioural trait, exploration:

Exploration

```
lmer_e <- lmer(exploration ~ scale(assay_rep, scale=FALSE) +
                 scale(body_size) +
                 (1|ID),
               data = df_syndrome)
```

```
rep_expl <- tidy(lmer_e, effects = "ran_pars", scales = "vcov") %>%
  select(group, estimate) %>%
  spread(group, estimate) %>%
  mutate(repeatability = ID/(ID + Residual))
```

ID	Residual	repeatability
0.362	0.909	0.285

Both of our traits of interest are repeatable at the among-individual level — the remaining question is characterising the association between these personality traits. Are individuals that are consistently bolder than average also more exploratory than average (and vice versa)?

Correlation using BLUPs

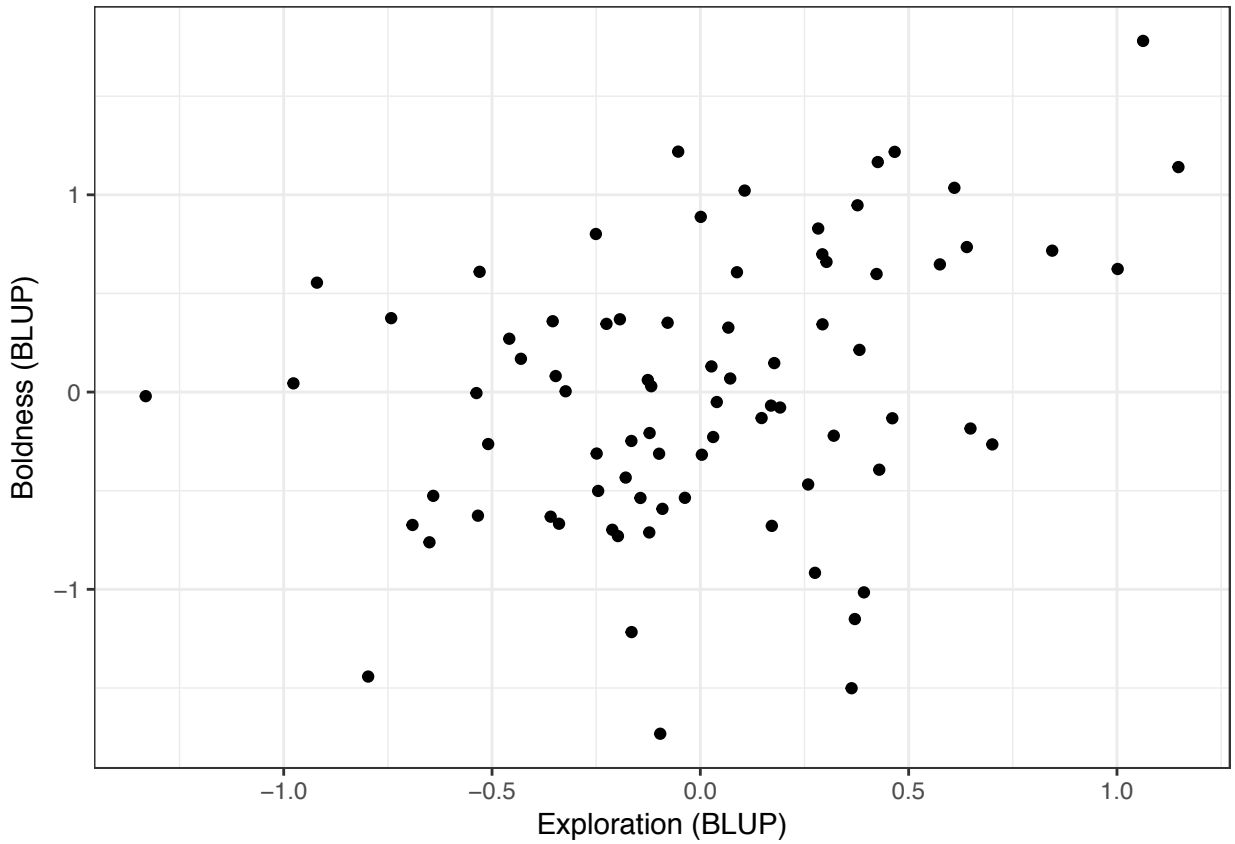
In our paper, we advise against the use of BLUPs due to their potential for spurious results due to anticonservative hypothesis tests and/or confidence intervals.

Here we will run through this method, purely so that we can then contrast the results with those that we get having (correctly) estimated the among-individual correlation between these behaviours directly from a multivariate model (in this case, bivariate).

We create two data frames of individual predictions extracted from model fits, one for each of our univariate `lme4` models for boldness and exploration. We then join these (by individual ID) to create a single data frame:

```
df_BLUPS_B <- data_frame(ID = row.names(ranef(lmer_b)$ID),
                         BLUP_B = ranef(lmer_b)$ID[,"(Intercept)"])

df_BLUPS_E <- data_frame(ID = row.names(ranef(lmer_e)$ID),
                         BLUP_E = ranef(lmer_e)$ID[,"(Intercept)"])

df_BLUPS_EB <- left_join(df_BLUPS_E,
                         df_BLUPS_B,
                         by = "ID")
```

We can plot these to see what our expectation of a correlation might be:



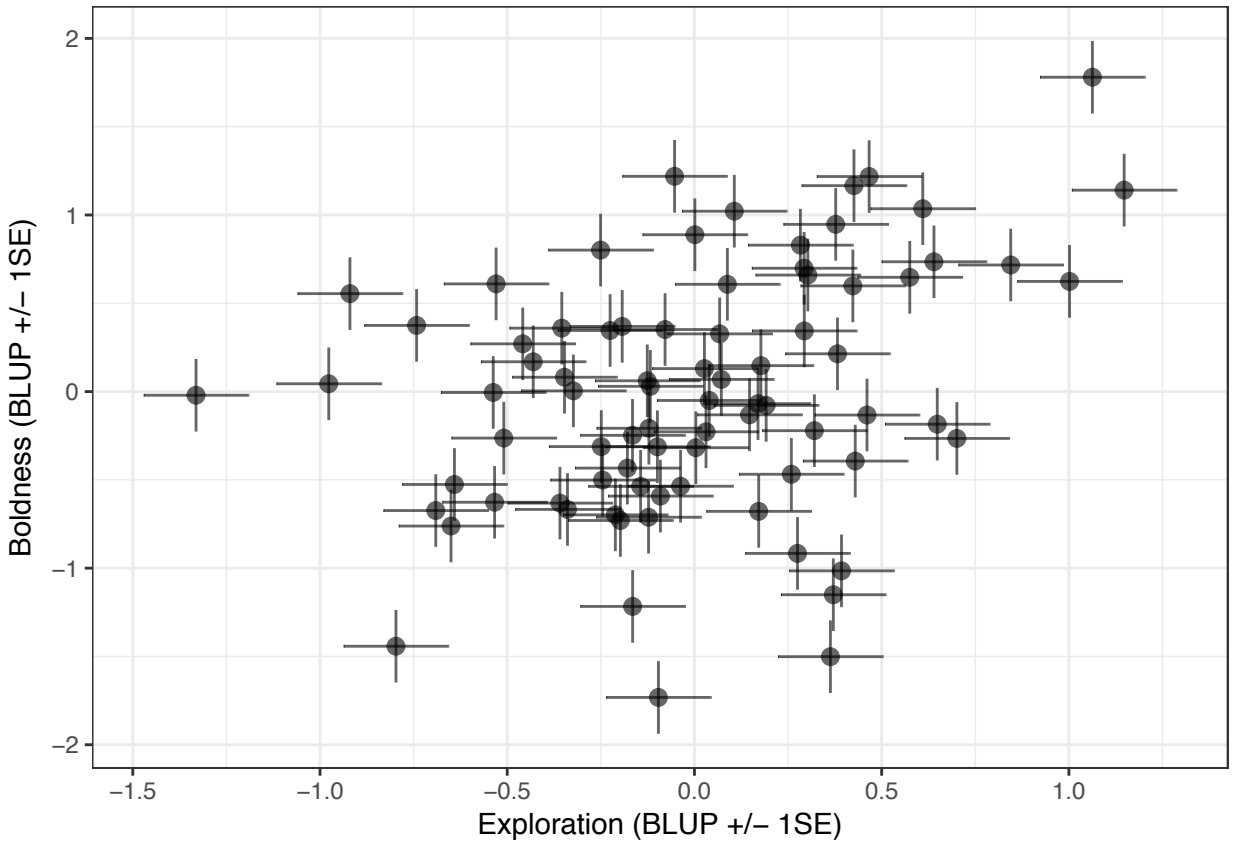
..and then simply perform a correlation test of these two traits using the `cor.test` function:

```
cor.test(df_BLUPS_EB$BLUP_E,
         df_BLUPS_EB$BLUP_B)
```

```
##
##  Pearson's product-moment correlation
##
## data:  df_BLUPS_EB$BLUP_E and df_BLUPS_EB$BLUP_B
## t = 3.2131, df = 78, p-value = 0.00191
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1320924 0.5223645
## sample estimates:
##       cor
## 0.3418867
```

As you can see, we get a positive correlation with a very small p-value ($P = 0.0019$), indicating that these traits are involved in a behavioural syndrome. While the correlation itself is fairly weak ($r = 0.34$), it appears to be highly significant, and suggests that individuals that are bolder than average also tend to be more exploratory than average.

However, as discussed in our paper (and in greater detail by Hadfield *et al*), using BLUPs in this way leads to anticonservative significance tests. This is because the error inherent in their prediction is not carried forward from the `lmer` models to the subsequent analysis (in this case, a correlation test). To illustrate this point quickly, below we plot the individual estimates along with their associated standard errors:



We now go on to estimate the correlation between these behaviours directly in a multivariate model, using `ASreml`.

Bivariate models

The correct approach for testing the hypothesised behavioural syndrome uses both response variables in a two-trait ('bivariate') mixed model. This model estimates the among-individual variance for each response variable (and the covariance between them). Separate (co)variances are also fitted for the residual variation. The bivariate model also allows for fixed effects to be fitted on both response variables.

We set up our model using the `asreml` function call, with our bivariate response variable being **exploration** and **boldness** bound together using `cbind`. You will also note that we `scale` our response variables, meaning that each is centred at their mean value and standardised to units of 1 standard deviation. This is not essential, but simply makes it easier for the model to be fit. Scaling the response variables also aids our understanding of the output, as both boldness and exploration are now on the same scale.

```
asr_E_B_us <- asreml(cbind(scale(exploration),
                           scale(boldness)) ~ trait +
                       trait:scale(assay_rep, scale = FALSE) +
                       trait:scale(body_size),
                     random =~ ID:us(trait, init = c(1,
                                                     0.1,1)),
                     rcov =~ units:us(trait, init = c(0.1,
                                                      0.1,0.1)),
                     data = df_syndrome,
                     maxiter = 100)
```

On the right hand side of our model formula, we use the `trait` keyword to specify that this is a multivariate model — `trait` itself tells the model to give us the intercept for each trait. We then interact `trait` with our fixed effects, **assay_rep** and **body_size**, so that we get estimates for the effect of these variables on each of our behaviours.

Our random effects structure starts with the `random` effects, where we tell the model to fit an 'unstructured' (`us`) covariance matrix for the grouping variable **ID**. This means that we want to calculate the variance in exploration due to differences among individuals, the variance in boldness due to differences among individuals, and the **covariance** between these variances.

Next, we set a structure for the residual variation (`rcov`), which is also sometimes known as the 'within-individual variation'. As we have repeated measures for both traits at the individual level, we also set an unstructured covariance matrix, which finds the residual variance for each trait and also allows the residuals to covary across the two traits.

Finally, we provide the name of the data frame, and a maximum number of iterations for `ASReml` to attempt to fit the model.

After the model has been fit by `ASReml`, we can check the fit using the same type of model diagnostic plots as we use for `lme4`:

```
plot(residuals(asr_E_B_us)~fitted(asr_E_B_us))
qqnorm(residuals(asr_E_B_us))
hist(residuals(asr_E_B_us))
```

The `summary` part of the `ASReml` model fit contains a large amount of information, so it is best to look only at certain parts of it at a single time. While we are not particularly interested in the fixed effects for current purposes, you can inspect these using the following code to check whether there were any large effects of assay repeat or body size on either trait:

```
summary(asr_E_B_us, all=T)$coef.fixed
```

We can see that there is a separate intercept for both personality traits (no surprise that these are very close to zero, given that we mean-centred and scaled each trait before fitting the model), and an estimate of the effect of assay repeat and body size on both traits. None of these appear to be large effects, so let's move on to the more interesting parts — the random effects estimates:

```
summary(asr_E_B_us)$varcomp
```

```
##                                          gamma component  std.error
## ID:trait!trait.exploration:exploration 0.2863101 0.2863101 0.07637361
## ID:trait!trait.boldness:exploration    0.0883864 0.0883864 0.06067166
## ID:trait!trait.boldness:boldness       0.3733306 0.3733306 0.08607573
## R!variance                             1.0000000 1.0000000         NA
## R!trait.exploration:exploration        0.7184419 0.7184419 0.06572786
## R!trait.boldness:exploration           0.3263211 0.3263211 0.04829180
## R!trait.boldness:boldness              0.6274169 0.6274169 0.05740290
##                                          z.ratio constraint
## ID:trait!trait.exploration:exploration  3.748810   Positive
## ID:trait!trait.boldness:exploration     1.456799   Positive
## ID:trait!trait.boldness:boldness        4.337234   Positive
## R!variance                                    NA      Fixed
## R!trait.exploration:exploration        10.930554   Positive
## R!trait.boldness:exploration            6.757279   Positive
## R!trait.boldness:boldness              10.930055   Positive
```

In the above summary table, we have the among-individual (co)variances listed first (starting with `ID`), then the residual (or within-individual) (co)variances (starting with `R`). You will notice that the variance estimates here are actually close to the `lme4` repeatability estimates, because our response variables were scaled to phenotypic standard deviations. We can also find the 'adjusted repeatability' (i.e., the repeatability conditional on the fixed effects) for each trait by dividing its among-individual variance estimate by the sum of its among-individual and residual variances.

Here, we use the `pin` function from the `nadiv` package (Wolak 2012) to estimate the repeatability and its standard error for each trait, conditional on the effects of assay repeat and body size. For this function, we provide the name of the model object, followed by a name that we want to give the estimate being returned, and a formula for the calculation. Each 'V' term in the formula refers to a variance component, using its position in the model summary shown above.

```
nadiv:::pin(asr_E_B_us, prop_expl ~ V1/(V1+V5))
nadiv:::pin(asr_E_B_us, prop_bold ~ V3/(V3+V7))
```

```
##           Estimate         SE
## prop_expl 0.284956 0.06113612
##            Estimate         SE
## prop_bold 0.3730518 0.06124283
```

We can also use this function to calculate the estimate and standard error of the correlation from our model (co)variances. We do this by specifying the formula for the correlation:

```
nadiv:::pin(asr_E_B_us, cor ~ V2/(sqrt(V1)*sqrt(V3)))
```

```
##      Estimate        SE
## cor 0.2703462 0.1594158
```

In this case, the estimate is similar (here, slightly lower) than our correlation estimate using BLUPs. However, if we consider confidence intervals as +/- 1.96SE around the estimate, the lower bound of the confidence interval would actually be -0.042. **With confidence intervals straddling zero, we would conclude that this correlation is likely non-significant**. As the use of standard errors in this way is only approximate, we should also test our hypothesis formally using likelihood ratio tests.

Hypothesis testing

We can now test the statistical significance of this correlation directly, by fitting a second model without the among-individual covariance between our two behavioural traits, and then using a likelihood ratio test to determine whether the model with the covariance produces a better fit.

Here, we use the `idh` structure for our random effects. This stands for 'identity matrix' (i.e., with 0s on the off-diagonals) with heterogeneous variances (i.e., the variance components for our two response traits are allowed to be different from one another). The rest of the model is identical to the `us` version.

```
asr_E_B_idh <- asreml(cbind(scale(exploration),
                            scale(boldness)) ~ trait +
                        trait:scale(assay_rep, scale = FALSE) +
                        trait:scale(body_size),
                      random =~ ID:idh(trait, init = c(1,1)),
                      rcov =~ units:us(trait, init = c(0.1,
                                                       0.1,0.1)),
                      data = df_syndrome,
                      maxiter = 100)
```

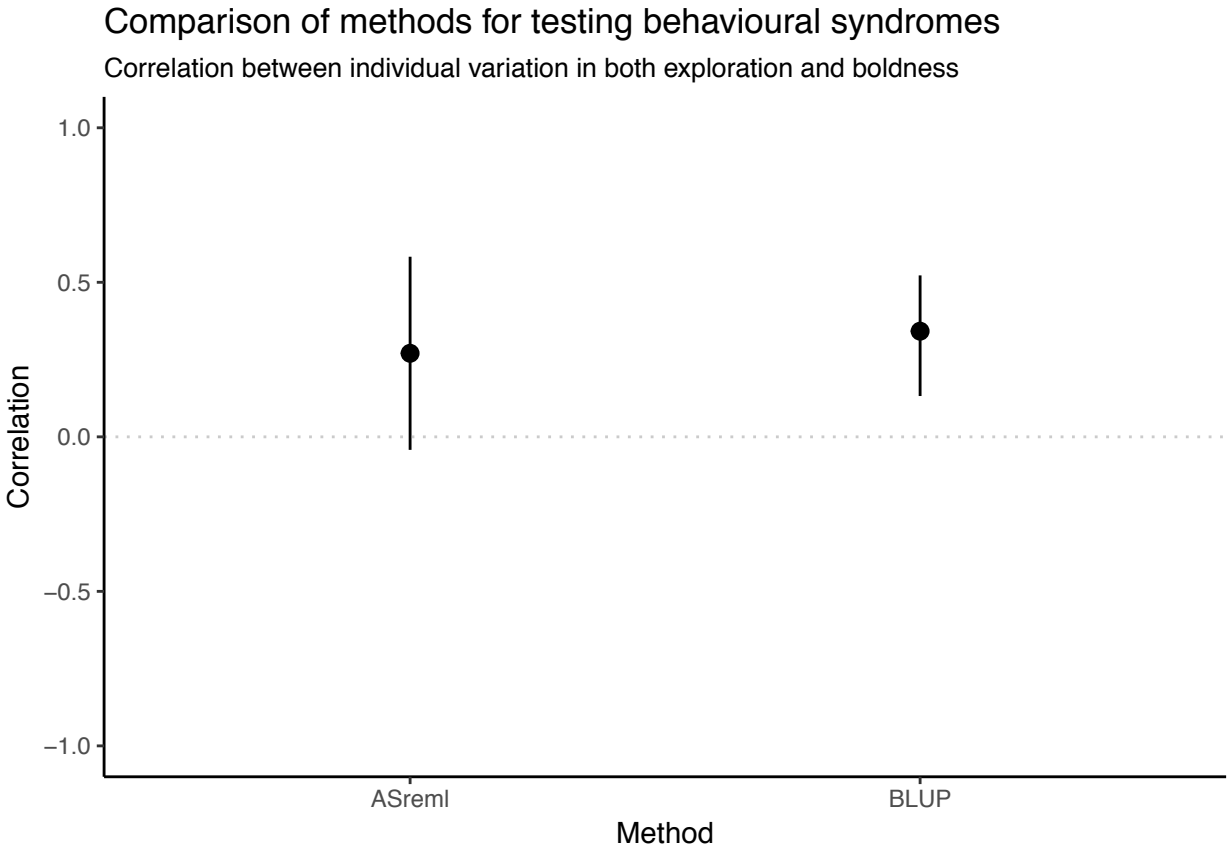
The likelihood ratio test is calculated as twice the difference between model log-likelihoods, on a single degree of freedom (the covariance term):

```
pchisq(2*(asr_E_B_us$loglik - asr_E_B_idh$loglik),
           1, lower.tail = FALSE)
```

```
## [1] 0.1170403
```

In sharp contrast to the highly-significant P-value given by a correlation test using BLUPs, here we find **no evidence for a behavioural syndrome between exploration and boldness**.

To better understand why BLUPs produce an anticonservative p-value in comparison to multivariate models, we should plot the correlation estimates and their confidence intervals. The confidence intervals are taken directly from the `cor.test` function for BLUPs, and for `ASReml` they are calculated as 1.96 times the standard error from the `pin` function.



Here we can clearly see that the BLUPs method - having failed to carry through the error around the predictions of individual-level estimates - is anticonservative, with small confidence intervals and a correspondingly small P-value ($P = 0.0019$). Testing the syndrome directly in a bivariate model that retains all the data, by comparison, enables us to capture the true uncertainty about the estimate of the correlation. This is reflected in the larger confidence intervals and, in this case, the non-significant P-value ($P = 0.117$).

Adding further traits

As part of our data collection, we also have a single value of mating success for each individual (which we will use as a proxy for fitness). We are interested in whether our personality traits are associated with variation

in this fitness-related measure. While our test above showed that the correlation between the measured personality traits was not significant, there did appear to be some relationship — so we shall incorporate both personality traits and fitness into a single trivariate model for hypothesis testing.

In this case, because the new response variable to be added to our model is fitness, we are **not** going to mean-centre and scale by phenotypic standard deviations, but instead divide by the mean fitness value (such that we are investigating among-individual covariance between personality traits and **relative fitness**). We create this new variable, `rel_fitness`, as follows:

```
df_syndrome <- df_syndrome %>%
  mutate(rel_fitness = fitness/mean(fitness, na.rm=TRUE))
```

Note that we will refer to this relative fitness trait simply as 'fitness' below for simplicity's sake.

Setting up the model

Below, we will set up our main model, which will allow for heterogeneous among-individual variances in our 3 traits (boldness, exploration, fitness), and will estimate the associations between them. Note, however, that we will use the `corgh` structure instead of `us` in the random effects. These structures fit the same model, but on a correlation rather than covariance scale. Note in this case we are just using `corgh` because it makes it easier in `ASReml` to specify some constraints that we require and (as we will see later, we can always backcalculate the covariances from the estimated correlations if we want them).

First, we set up starting values from the model, which we also use to set some constraints. We set constraints in `ASReml` by specifying some starting values in a numeric vector, then giving each value a 'name' that corresponds to how `ASReml` should treat the corresponding part of the random effects matrix during model fitting:

- U: Unconstrained (can take any value, positive or negative)
- P: Positive (must be a positive value)
- F: Fixed (remains fixed at the given value)

An important point: while the starting values (`init`) for the `us` structure were provided in the form of the lower triangle of a covariance matrix, for `corgh` we provide the correlations first, and then the variances.

For the random effects, we set generic starting values — the 3 correlations have starting values close to 0 and are unconstrained, while the variance components have starting values of unit variance (and are constrained to be positive values):

```
init_E_B_fit_cor <- c(0.1,
                      0.1,0.1,
                      1,1,1)
names(init_E_B_fit_cor) <- c("U",
                             "U","U",
                             "P","P","P")
```

For the residuals (or 'within-individual' variance), we must bear in mind that we have only a single fitness value per individual — therefore, that trait has **no within-individual variance**, and **within-individual correlations involving fitness must be set to zero as they cannot be estimated**. We set the starting value for both correlations to 0 below, and denote them as fixed at those values using 'F'. The variance component is slightly trickier — variances have to be positive, therefore we simply fix the within-individual variance at a very small positive number (here, 1e-08 — i.e., so small as to be effectively 0):

```
init_E_B_fit_res <- c(0.1,
                      0,0,
                      0.1, 0.1, 1e-08)
names(init_E_B_fit_res) <- c("U",
                             "F","F",
                             "P","P","F")
```

Now, we can fit our model with these starting values and constraints. Again, we `cbind` our response variables on the left-hand side of the formula, and use `trait` to denote a multivariate model. Remember that we have created the 'relative fitness' variable by essentially scaling by its mean, so this does not need to be scaled as the behavioural traits are.

We can also use the `at` keyword to specify that fixed effects are estimated only for certain traits — here, we test for an effect of assay repeat only on exploration and boldness (because these were measured repeatedly), while we test for the effect of body size on all of our traits.

Fit the model as follows (and be sure to use visual diagnostic checks of the residuals):

```
asr_E_B_fit_cor <- asreml(cbind(scale(exploration),
                                scale(boldness),
                                rel_fitness) ~ trait +
                            at(trait,1):assay_rep +
                            at(trait,2):assay_rep +
                            trait:scale(body_size),
                          random =~ ID:corgh(trait, init = init_E_B_fit_cor),
                          rcov =~ units:corgh(trait, init = init_E_B_fit_res),
                          data = df_syndrome,
                          maxiter = 500)
```

We can take a quick look at the fixed effects:

```
summary(asr_E_B_fit_cor, all=T)$coef.fixed
```

Below, we specify that we want to look at the variance components using `$varcomp`. In the interests of space, we will request only the `component` (i.e., the variance estimate) and its `std.error`:

```
summary(asr_E_B_fit_cor)$varcomp[,c("component","std.error")]
```

```
##                                                  component   std.error
## ID:trait!trait.boldness:!trait.exploration.cor    0.27031497 0.159419988
## ID:trait!trait.rel_fitness:!trait.exploration.cor 0.23386699 0.138687881
## ID:trait!trait.rel_fitness:!trait.boldness.cor    0.66168293 0.087961997
## ID:trait!trait.exploration                        0.28630613 0.076372770
## ID:trait!trait.boldness                           0.37322016 0.086051330
## ID:trait!trait.rel_fitness                        0.05659086 0.009060437
## R!variance                                        1.00000000          NA
## R!trait.boldness:!trait.exploration.cor           0.48603894 0.049410253
## R!trait.rel_fitness:!trait.exploration.cor        0.00000000          NA
## R!trait.rel_fitness:!trait.boldness.cor           0.00000000          NA
## R!trait.exploration                               0.71844420 0.065728071
## R!trait.boldness                                  0.62744922 0.057405898
## R!trait.rel_fitness                               0.00000001          NA
```

Here we can see that the fit provides us with estimates and standard errors of:

- 3 among-individual correlations;
- 3 among-individual variance components;
- 3 within-individual correlations;
- 3 within-individual variance components.

You can see from the estimates that our constraints have worked in the model: within-individual correlations featuring fitness are at 0, and the residual fitness variance is a very small positive number (such that all the variation is at the among-individual level).

A quick sanity check also tells us that the correlation between boldness and exploration (the first variance component in our summary table above, $r = 0.27$ SE 0.159) estimated in this model is the same as in our earlier bivariate model.

From a first glance at the correlation estimates and their associated standard errors, it appears likely that there is a significant among-individual correlation between relative fitness and boldness ($r = 0.662$ SE 0.088), but not between relative fitness and exploration ($r = 0.234$ SE 0.139).

Hypothesis testing

We can again use likelihood ratio tests for hypothesis testing with these models. We first test for an association between relative fitness and our bivariate personality phenotype (defined by the two traits). We do this by fixing both correlations with fitness ($r_{\text{boldness,fitness}}$ and $r_{\text{exploration,fitness}}$) to 0. We then use a likelihood ratio test to analytically compare our main model (with all correlations estimated) to this second model (no correlation between fitness and boldness/exploration), which tests whether allowing those correlations provides a statistically significant improvement in the model fit. Note this is not testing the significance of each trait-fitness correlation separately, it is testing whether there is any significant fitness-phenotype correlation overall.

We set the correlations to 0 as follows:

```
init_E_B_fit_cor_FEB0 <- c(0.1,
                           0,0,
                           1,1,1)
names(init_E_B_fit_cor_FEB0) <- c("U",
                                  "F","F",
                                  "P","P","P")

asr_E_B_fit_cor_FEB0 <- asreml(cbind(scale(exploration),
                                     scale(boldness),
                                     rel_fitness) ~ trait +
                                 at(trait,1):assay_rep +
                                 at(trait,2):assay_rep +
                                 trait:scale(body_size),
                               random =~ ID:corgh(trait, init = init_E_B_fit_cor_FEB0),
                               rcov =~ units:corgh(trait, init = init_E_B_fit_res),
                               data = df_syndrome,
                               maxiter = 800)
```

We then test the difference in model fits using a likelihood ratio test with 2 degrees of freedom:

```
pchisq(2*(asr_E_B_fit_cor$loglik - asr_E_B_fit_cor_FEB0$loglik),
           2, lower.tail = FALSE)
```

```
## [1] 5.654352e-07
```

Here we find evidence of significant correlation structure — based on the estimates and SEs from the model summary, it's a fairly safe bet that this is being driven by the fitness-boldness association. If tests of each of the specfic trait-fitness correlations are needed, we advise using pairwise models (but note of course that multiple testing issues might require consideration if you want to statistically test every pairwise correlation estimate and you have a lot of traits). We will fit the two bivariate trait-fitness models below for completeness, and they should confirm our suspicions about which personality trait is driving the correlation between the bivariate behavioural phenotype and fitness.

As with tests of the earlier bivariate models for behavioural syndromes, we fit models with both `us` and `idh` structures (or `corgh` with setting the correlation to 0) for hypothesis testing using likelihood ratio tests. In this case, we also have to set the residual variation in fitness to a very small (near-zero) positive number, and we do not fit a residual covariance. Here we demonstrate for boldness and fitness:

```
init_fitbiv_res <- c(0.1,1e-08)
names(init_fitbiv_res) <- c("P","F")

asr_B_fit_us <- asreml(cbind(scale(boldness),
                             rel_fitness) ~ trait +
                         at(trait,1):assay_rep +
                         trait:scale(body_size),
                       random =~ ID:us(trait, init = c(1,
                                                       0.1,1)),
                       rcov =~ units:idh(trait, init = init_fitbiv_res),
                       data = df_syndrome,
                       maxiter = 800)

asr_B_fit_idh <- asreml(cbind(scale(boldness),
                              rel_fitness) ~ trait +
                          at(trait,1):assay_rep +
                          trait:scale(body_size),
                        random =~ ID:idh(trait, init = c(1,1)),
                        rcov =~ units:idh(trait, init = init_fitbiv_res),
                        data = df_syndrome,
                        maxiter = 800)
```

```
## [1] 8.164003e-08
```

We can now run the same test for exploration and fitness:

```
## [1] 0.1024701
```

As we had anticipated from the estimate and standard error of the correlations in our trivariate model, the association between individual variation in boldness and relative fitness is significant, while there is no evidence for a significant association between individual variation in exploration and fitness.

A slight digression: converting correlations back to covariances can be useful

While we set up the trivariate model to output results in terms of correlation matrices, we could have fit the model on a covariance scale using `us`. While correlations are intuitive, sometimes having the answers on the covariance scale is useful. For instance, in the current example, the trait-fitness correlations could be used to infer selection — but if we wanted to express the strength of that selection, the normal way to do so is through ___selection differentials. These are the trait – (relative) fitness covariances, and/or selection gradients (the partial regressions of relative fitness on traits which can be calculated from variance and covariance terms).

Since a correlation is simply the covariance rescaled by the product of the squared variances, we can retrieve the covariance terms by simply rearranging as follows:

$$COV_{T1,T2} = r_{T1,T2} \times \sqrt{V_{T1}} \times \sqrt{V_{T2}}$$

Again, the `pin` function comes to our rescue. As an example, we can get the covariance between exploration and boldness from our trivariate model (with `corgh` correlation-structure) as follows:

```
nadiv:::pin(asr_E_B_fit_cor, cov_E_B ~ V1*sqrt(V4)*sqrt(V5))
```

```
##           Estimate         SE
## cov_E_B 0.08836249 0.06066255
```

We might want to present our final results as a matrix with variances on the diagonals, covariances below and correlations above (with standard errors in parentheses):

	Exploration	Boldness	Fitness
Exploration	0.29 (0.08)	0.27 (0.16)	0.23 (0.14)
Boldness	0.09 (0.06)	0.37 (0.09)	0.66 (0.09)
Fitness	0.03 (0.02)	0.1 (0.02)	0.06 (0.01)

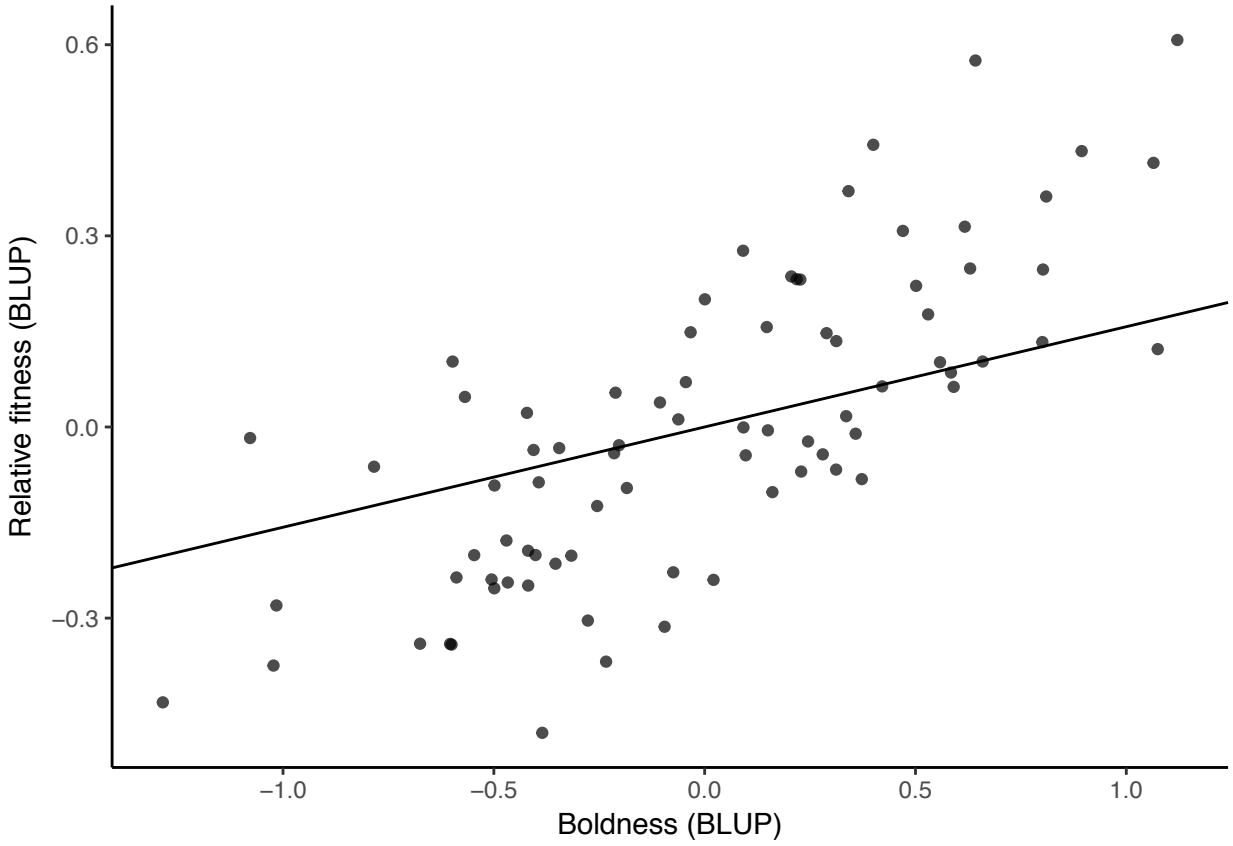
Conclusions

To conclude, then: we found that the correlation between boldness and exploration tends to be positive among male haggis. This correlation is not statistically significant, and thus does not provide strong evidence for a behavioural syndrome. However, inappropriate analysis of BLUP extracted from univariate models would lead to a different (erroneous) conclusion. We also found no statistically significant association between among-individual variation in exploration and fitness. However, we did find a statistically significant positive association between among-individual variation in boldness and our fitness proxy, indicating that bolder male haggis had greater mating success (see figure below).

Note: below, we use BLUPs from our trivariate model to construct a figure that illustrates the association between boldness and fitness. Unlike its use in secondary statistical analyses, this is an appropriate use of BLUPs — i.e., just for illustrative purposes!

```
# Retrieve BLUPs from ASReml trivariate model
#  and reform into data frame for plotting
df_bf_coefs <- data_frame(Trait = attr(asr_E_B_fit_cor$coefficients$random, "names"),
                          Value = asr_E_B_fit_cor$coefficients$random) %>%
  separate(Trait, c("ID","Trait"), sep = ":") %>%
  filter(Trait %in% c("trait_boldness", "trait_rel_fitness")) %>%
  spread(Trait, Value)
```

```
# Find the regression line -
#  the covariance of boldness, relative fitness divided by
#  the square root of the variance in boldness
B_fit_slope <-  as.numeric(nadiv:::pin(asr_E_B_fit_cor,
                                       slope ~ (V3*sqrt(V5)*sqrt(V6))/
                                         sqrt(V5))$Estimate)

ggplot(df_bf_coefs, aes(x = trait_boldness, y = trait_rel_fitness, group = ID)) +
  geom_point(alpha = 0.7) +
  geom_abline(intercept = 0, slope = B_fit_slope) +
  labs(x = "Boldness (BLUP)",
       y = "Relative fitness (BLUP)") +
  theme_classic()
```



Further tutorials

We will continue to develop tutorials for multivariate modelling of individual (co)variation, which will cover some of the more advanced issues discussed in our paper. Please visit tomhouslay.com for more information.

Avoiding the misuse of BLUP in behavioral ecology: Multivariate modelling for individual variation (MCMCglmm tutorial)

T.M. Houslay & A.J. Wilson, Behavioral Ecology

January 2017

Introduction

Overview

This tutorial accompanies our paper, "Avoiding the misuse of BLUP in behavioral ecology". Below, we provide worked examples of multivariate statistical methods for directly testing hypotheses about associations between individual variation in behaviour and other traits. Below, we will:

- Test the correlation between two personality traits (behaviours measured repeatedly on individuals);
- Test for an association between these personality traits and a measure of fitness (one value per individual).

In this version of the tutorial, we illustrate these models using the `R` package `MCMCglmm`, developed by Jarrod Hadfield. Visit the CRAN page for `MCMCglmm` here for links and citation info: https://cran.r-project.org/web/packages/MCMCglmm/index.html.

MCMCglmm fits generalised linear mixed modes (GLMMs) in a Bayesian framework, using Markov chain Monte Carlo techniques. We have also provided a separate tutorial for the `R` interface for `ASReml`, which fits GLMMs using maximum likelihood (and so is likely more familiar to `lme4` users) but is commercially licensed software.

Aims

Please note that we do assume readers are familiar with the general principles of specifying mixed effects models, and in particular with the use of `MCMCglmm` for univariate mixed effects models. Readers unfamiliar with using univariate mixed effects models for modelling a single behavioural trait might prefer to start with Dingemanse & Dochtermann's 2013 paper, 'Quantifying individual variation in behaviour: mixed effects modelling approaches'. Readers unfamiliar with `MCMCglmm` should look at Jarrod Hadfield's excellent course notes, available at the `MCMCglmm` CRAN page: https://cran.r-project.org/web/packages/MCMCglmm/index.html.

We also use various methods for manipulating and visualising data frames using the `tidyverse` package (including `tidyr`, `dplyr`, `ggplot2` etc) — more details on their use can be found at http://r4ds.had.co.nz/.

In our tutorial, we aim to teach the following:

- How to phrase questions of interest in terms of variances and covariances (or derived correlations or regressions);
- How to incorporate more advanced model structures, such as:
- Fixed effects that apply only to a subset of the response traits;
- Traits which are measured a different number of times (*e.g.*, repeated measures of behaviour and a single value of breeding success);
- Interpreting MCMC credible intervals.

Packages required

There are several packages that you must have installed in `R` prior to starting this tutorial:

- `MCMCglmm`
- `lme4`
- `nadiv`
- `tidyverse`
- `broom`

'Study system'

For this tutorial, we have collected data on populations of wild haggis that roam the Highlands of Scotland.

Figure 1: A male haggis in the wild (thanks to Emma Wood, http://www.ewood-art.co.uk/)

We tag all haggis individually when they emerge from their burrows as juveniles in their first spring. Here, we concentrate on male haggis, which are solitary and territorial. Previous work has identified behaviours that can be measured repeatedly, and used to represent three personality traits: **boldness**, **exploration**, and **aggression**. We also have the ability to collect a single measure of mating success (as a fitness proxy) for each male at the end of the season.

Behavioural syndromes

One type of 'behavioural syndrome' is a correlation between personality traits. Since personality can be viewed (under most definitions) as the repeatable (among-individual) component of behaviour, evidence for the presence of a behavioural syndrome is provided by covariance among behaviours that arises from among-individual differences.

Here we have repeatedly measured behaviours that represent **boldness** and **exploration**. We observed each behaviour 4 times per individual. We also measured their body size on the day of behavioural assay to control for general size effects. in our statistical models.

Load libraries and inspect data

```
library(lme4)
library(MCMCglmm)
library(tidyverse)
library(broom)
library(nadiv)

df_syndrome <- read_csv("syndrome.csv")
```

This data frame has 6 variables:

- Individual **ID**
- The repeat number for each behavioural test, **assay_rep**
- **boldness**, measured 4 times per individual
- **exploration**, measured 4 times per individual
- **fitness**, our measure of mating success, with a single value for each individual
- Individual **body_size**, as measured on the day of testing.

Univariate models

We first use the `R` package `lme4` to determine the proportion of phenotypic variation (adjusted for fixed effects) that is due to differences among individuals, separately for each behaviour. We assume readers have knowledge of these 'univariate' models and their use in behavioural studies — if not, there are various other publications that go into them in greater detail (e.g., Dingemanse & Dochtermann (2013)).

Boldness

Our model includes fixed effects of the assay repeat number (centred) and individual body size (centred and scaled to standard deviation units), as we wish to control for any systematic effects of these variables on individual behaviour. Please be aware that controlling variables are at your discretion — for example, while we want to characterise among-individual variance in boldness after controlling for size effects in this study, others may wish to characterise among-individual variance in boldness without such control. Indeed, using the techniques shown later in this tutorial, it would be entirely possible to characterise both among-individual variance in boldness and in size, and the among-individual covariance between these measurements.

```
lmer_b <- lmer(boldness ~ scale(assay_rep, scale=FALSE) +
                 scale(body_size) +
                 (1|ID),
               data = df_syndrome)
plot(lmer_b)
qqnorm(residuals(lmer_b))
hist(residuals(lmer_b))

summary(lmer_b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: boldness ~ scale(assay_rep, scale = FALSE) + scale(body_size) +
##     (1 | ID)
##    Data: df_syndrome
##
## REML criterion at convergence: 1061.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.3645 -0.6496 -0.1154  0.6463  2.6894
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.6951   0.8337
##  Residual             1.1682   1.0808
## Number of obs: 320, groups:  ID, 80
##
## Fixed effects:
##                                 Estimate Std. Error t value
## (Intercept)                     20.09133    0.11108  180.87
## scale(assay_rep, scale = FALSE) -0.04805    0.05404   -0.89
## scale(body_size)                 0.14128    0.10893    1.30
##
## Correlation of Fixed Effects:
##             (Intr) s(_s=F
## s(_,s=FALSE  0.000
## scl(bdy_sz)  0.000 -0.002
```

Having examined diagnostic plots of the model fit, we can check the model `summary`. We are interested in the *random effects* section of the `lme4` model output (specifically the **variance** component — note that the standard deviation here is simply the square root of the variance). Evidence for 'animal personality' (or 'consistent among-individual differences in behaviour') in the literature is largely taken from the **repeatability** of behaviorual traits: we can compute this **repeatability** (also known as the *intraclass correlation coefficient*) by dividing the variance in the trait due to differences among individuals (V_{ID}) by the total phenotypic variance after accounting for the fixed effects ($V_{ID} + V_{residual}$). This can be done quickly and automatically through the use of the R package `broom`:

```
rep_bold <- tidy(lmer_b, effects = "ran_pars", scales = "vcov") %>%
  select(group, estimate) %>%
  spread(group, estimate) %>%
  mutate(repeatability = ID/(ID + Residual))

rep_bold
```

ID	Residual	repeatability
0.695	1.168	0.373

So we can see that 37.3% of the phenotypic variation in boldness (having controlled for body size and assay repeat number) is due to differences among individuals.

Let's do the same for our other behavioural trait, exploration:

Exploration

```
lmer_e <- lmer(exploration ~ scale(assay_rep, scale=FALSE) +
                 scale(body_size) +
                 (1|ID),
               data = df_syndrome)
```

```
rep_expl <- tidy(lmer_e, effects = "ran_pars", scales = "vcov") %>%
  select(group, estimate) %>%
  spread(group, estimate) %>%
  mutate(repeatability = ID/(ID + Residual))
```

ID	Residual	repeatability
0.362	0.909	0.285

Both of our traits of interest are repeatable at the among-individual level — the remaining question is characterising the association between these personality traits. Are individuals that are consistently bolder than average also more exploratory than average (and vice versa)?

Correlation using BLUPs

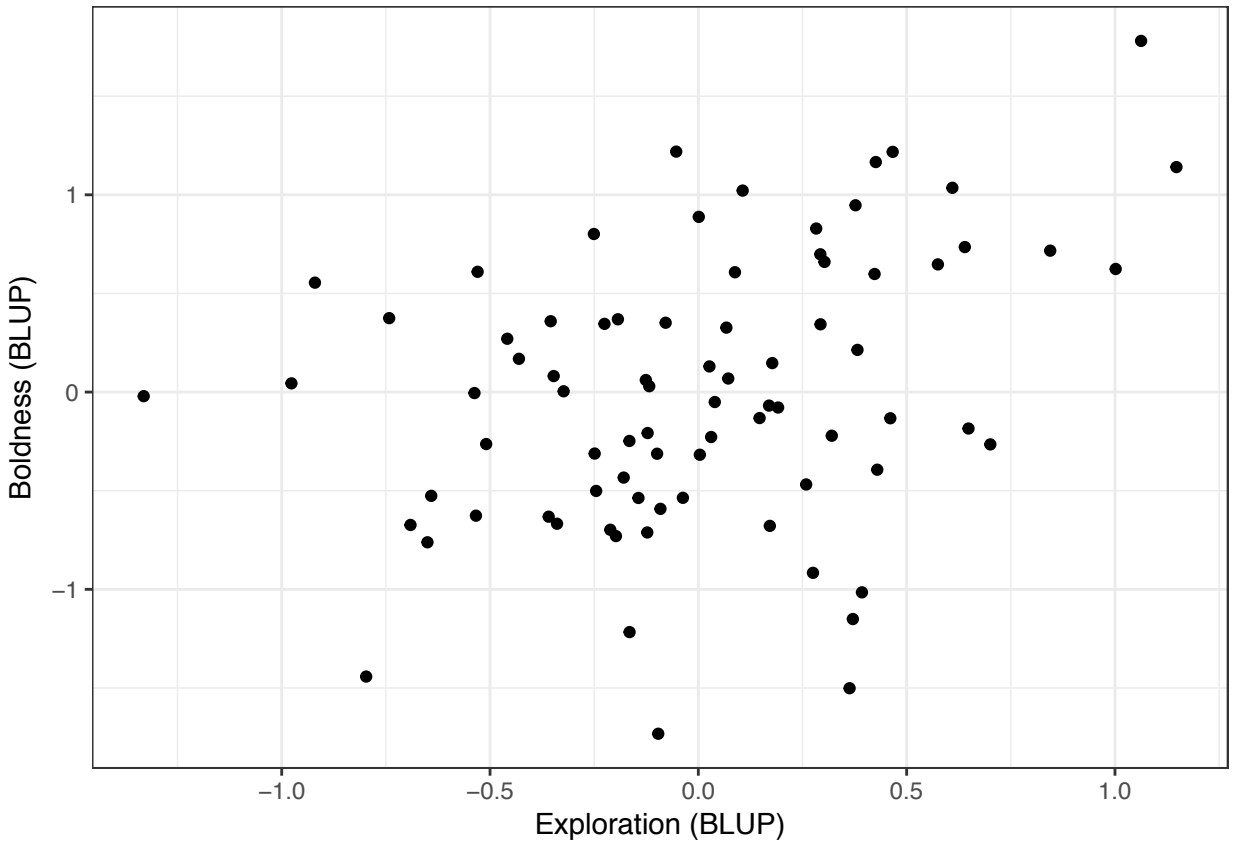
In our paper, we advise against the use of BLUPs due to their potential for spurious results due to anticonservative hypothesis tests and/or confidence intervals.

Here we will run through this method, purely so that we can then contrast the results with those that we get having (correctly) estimated the among-individual correlation between these behaviours directly from a multivariate model (in this case, bivariate).

We create two data frames of individual predictions extracted from model fits, one for each of our univariate `lme4` models for boldness and exploration. We then join these (by individual ID) to create a single data frame:

```
df_BLUPS_B <- data_frame(ID = row.names(ranef(lmer_b)$ID),
                         BLUP_B = ranef(lmer_b)$ID[,"(Intercept)"])

df_BLUPS_E <- data_frame(ID = row.names(ranef(lmer_e)$ID),
                         BLUP_E = ranef(lmer_e)$ID[,"(Intercept)"])

df_BLUPS_EB <- left_join(df_BLUPS_E,
                         df_BLUPS_B,
                         by = "ID")
```

We can plot these to see what our expectation of a correlation might be:



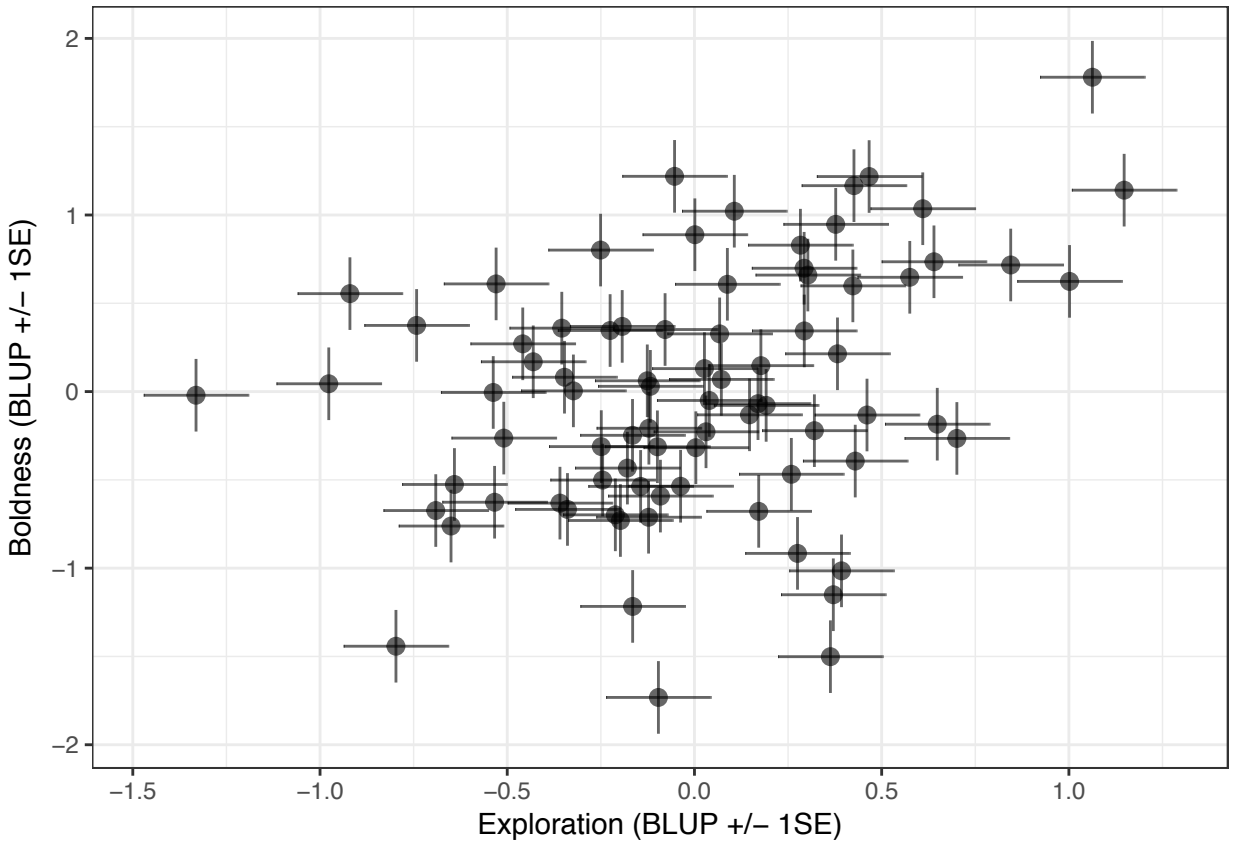
..and then simply perform a correlation test of these two traits using the `cor.test` function:

```
cor.test(df_BLUPS_EB$BLUP_E,
         df_BLUPS_EB$BLUP_B)
```

```
##
##  Pearson's product-moment correlation
##
## data:  df_BLUPS_EB$BLUP_E and df_BLUPS_EB$BLUP_B
## t = 3.2131, df = 78, p-value = 0.00191
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1320924 0.5223645
## sample estimates:
##       cor
## 0.3418867
```

As you can see, we get a positive correlation with a very small p-value ($P = 0.0019$), indicating that these traits are involved in a behavioural syndrome. While the correlation itself is fairly weak ($r = 0.34$), it appears to be highly significant, and suggests that individuals that are bolder than average also tend to be more exploratory than average.

However, as discussed in our paper (and in greater detail by Hadfield *et al*), using BLUPs in this way leads to anticonservative significance tests. This is because the error inherent in their prediction is not carried forward from the `lmer` models to the subsequent analysis (in this case, a correlation test). To illustrate this point quickly, below we plot the individual estimates along with their associated standard errors:



We now go on to estimate the correlation between these behaviours directly in a multivariate model, using `MCMCglmm`.

Bivariate models

The correct approach for testing the hypothesised behavioural syndrome uses both response variables in a two-trait ('bivariate') mixed model. This model estimates the among-individual variance for each response variable (and the covariance between them). Separate (co)variances are also fitted for the residual variation. The bivariate model also allows for fixed effects to be fitted on both response variables.

First, we need to create a 'prior' for our model. We recommend reading up on the use of priors; briefly, we use a parameter-expanded prior here that should be uninformative for our model. One of the model diagnostic steps that should be used later is to check that the model is robust to multiple prior specifications.

```
prior_E_B_1px = list(R = list(V = diag(2), nu = 0.002),
                     G = list(G1 = list(V = diag(2), nu = 2,
                                        alpha.mu = rep(0,2),
                                        alpha.V = diag(25^2,2,2))))
```

We set up our model using the `MCMCglmm` function call, with our bivariate response variable being **exploration** and **boldness** bound together using `cbind`. You will also note that we `scale` our response variables, meaning that each is centred at their mean value and standardised to units of 1 phenotypic standard deviation. This simply makes it easier for the model to be fit, and for us to understand the output, as both boldness and exploration are now on the same scale.

```
mcmc_E_B_us <- MCMCglmm(cbind(scale(exploration), scale(boldness)) ~ trait-1 +
                          trait:scale(assay_rep, scale = FALSE) +
                          trait:scale(body_size),
                        random =~  us(trait):ID,
                        rcov =~ us(trait):units,
                        family = c("gaussian","gaussian"),
                        prior = prior_E_B_1px,
                        nitt=420000,
                        burnin=20000,
                        thin=100,
                        verbose = TRUE,
                        data = as.data.frame(df_syndrome))
```

On the right hand side of our model formula, we use the `trait` keyword to specify that this is a multivariate model — `trait-1` effectively tells the model to give us a distinct intercept for each trait. We then interact `trait` with our fixed effects, **assay_rep** and **body_size**, so that we get estimates for the effect of these variables on each of our behaviours.

Our random effects structure starts with the `random` effects, where we tell the model to fit an 'unstructured' (`us`) covariance matrix for the grouping variable **ID**. This means that we want to calculate the variance in exploration due to differences among individuals, the variance in boldness due to differences among individuals, and the **covariance** between these variances.

Next, we set a structure for the residual variation (`rcov`), which is also sometimes known as the 'within-individual variation'. As we have repeated measures for both traits at the individual level, we also set an unstructured covariance matrix, which finds the residual variance for each trait and also allows these variances to covary.

We then provide the name of the object we set up as the model prior, and values for the total number of iterations (`nitt`), the 'burn-in' of initial iterations to be discarded as the model starts to converge (`burnin`), and the number of iterations to discard in between successive stored samplles (`thin`, which helps to reduce autocorrelation in sampling).

Finally, we provide the name of the data frame — we enclose this in the `as.data.frame` function as `MCMCglmm` does not work with the `tbl_df` format used in the **tidyverse** group of packages.

After the model has been fit by `MCMCglmm` (which will take some time!), we can check some model diagnostics using plots of the MCMC samples. Here we show just the plots for our variance components (these plots are also available for fixed effects, using `Sol`):

```
plot(mcmc_E_B_us$VCV)
```

For current purposes these should look fine, assuming you have used our simulated data and the settings above. Note however that for any real analysis various other tests (e.g. of autocorrelation, robustness to different priors, and good model convergence using the `geweke.diag` and `gelman.diag` diagnostic functions) should be used before accepting final results.

The `summary` part of the `MCMCglmm` model fit contains a large amount of information. Some general information at the start of the summary includes the model DIC. The `G-structure` then contains information about the random effects (co)variances, the `R-structure` the residual (co)variances, and the `Location effects` holds the fixed effects results information.

Each of these sections provides the mean of the posterior distribution returned by `MCMCglmm`, in addition to the lower and upper bounds of the 95% credible intervals. The effective sample size is also provided, and – for the fixed effects only – a pMCMC value.

```
summary(mcmc_E_B_us)
```

```
##
##  Iterations = 20001:419901
##  Thinning interval  = 100
##  Sample size  = 4000
##
##  DIC: 1596.616
##
##  G-structure:  ~us(trait):ID
##
##                                     post.mean l-95% CI u-95% CI eff.samp
## traitexploration:traitexploration.ID   0.29234  0.14609   0.4538     4000
## traitboldness:traitexploration.ID      0.08287 -0.03125   0.2079     4000
## traitexploration:traitboldness.ID      0.08287 -0.03125   0.2079     4000
## traitboldness:traitboldness.ID         0.38889  0.22405   0.5735     4000
##
##  R-structure:  ~us(trait):units
##
##                                       post.mean l-95% CI u-95% CI
## traitexploration:traitexploration.units    0.7340   0.5996   0.8697
## traitboldness:traitexploration.units       0.3338   0.2390   0.4353
## traitexploration:traitboldness.units       0.3338   0.2390   0.4353
## traitboldness:traitboldness.units          0.6391   0.5287   0.7614
##                                       eff.samp
## traitexploration:traitexploration.units     4000
## traitboldness:traitexploration.units        3365
## traitexploration:traitboldness.units        3365
## traitboldness:traitboldness.units           3685
##
##  Location effects: cbind(scale(exploration), scale(boldness)) ~ trait - 1 + trait:scale(assay_rep, s
##
##                                                  post.mean   l-95% CI
## traitexploration                                 0.0002371 -0.1503944
## traitboldness                                   -0.0013789 -0.1529724
## traitexploration:scale(assay_rep, scale = FALSE) -0.0226367 -0.1030113
## traitboldness:scale(assay_rep, scale = FALSE)    -0.0355084 -0.1083371
## traitexploration:scale(body_size)                 0.0714747 -0.0887465
## traitboldness:scale(body_size)                    0.1047925 -0.0543119
##                                                   u-95% CI eff.samp pMCMC
## traitexploration                                 0.1557892     4000 0.992
## traitboldness                                    0.1667160     4000 0.992
## traitexploration:scale(assay_rep, scale = FALSE) 0.0599347     4000 0.586
## traitboldness:scale(assay_rep, scale = FALSE)    0.0473711     4000 0.392
## traitexploration:scale(body_size)                0.2192468     3779 0.349
## traitboldness:scale(body_size)                   0.2627610     4000 0.184
```

Note that you will **not** have exactly the same results as we have, because of the way that the MCMC process works — if you run it again yourself, you will get slightly different answers again. However, they should be very similar.

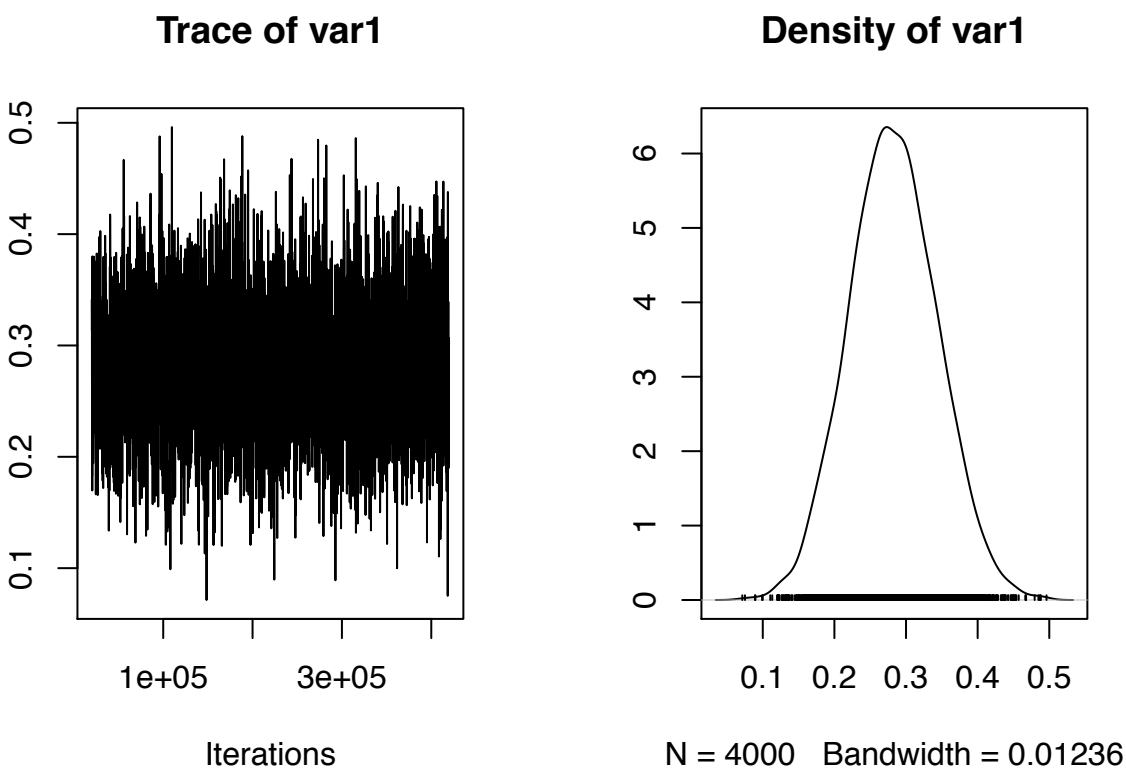
From the fixed effects, we can see that there is a separate intercept for both personality traits (no surprise that these are very close to zero, given that we mean-centred and scaled each trait before fitting the model),

and an estimate of the effect of assay repeat and body size on both traits. None of these appear to be large effects, so let's move on to the more interesting parts — the random effects estimates.

In the `G-structure`, we have the among-individual (co)variances. These are given such that they can be reformed into a matrix, which is why $V_{boldness}$ and $V_{exploration}$ are shown once each, while the among-individual covariance between them ($COV_{boldness,exploration}$) is shown twice.

You will notice that the variance estimates here are actually close to the `lme4` repeatability estimates, which is because we scaled our response variables to phenotypic standard deviations. We can also find the 'adjusted repeatability' (i.e., the repeatability conditional on the fixed effects) for each trait by dividing its among-individual variance estimate by the sum of its among-individual and residual variances. To do this, we can create a new posterior distribution of (for example) 'proportion of exploration variance explained by differences among individuals'. We do this by referencing the different variance components by their name as shown in the summary (note that sometimes different versions display these with or without the 'trait' prefix, so check how yours has displayed).

```
mcmc_prop_E <- mcmc_E_B_us$VCV[,"traitexploration:traitexploration.ID"]/(
  mcmc_E_B_us$VCV[,"traitexploration:traitexploration.ID"] +
    mcmc_E_B_us$VCV[,"traitexploration:traitexploration.units"]
)

plot(mcmc_prop_E)
```



We can interrogate this new distribution for its mean and 95% CIs:

```
mean(mcmc_prop_E)
```

```
## [1] 0.2824676
```

```
HPDinterval(mcmc_prop_E)
```

```
##          lower     upper
## var1 0.1620258 0.3991629
## attr(,"Probability")
## [1] 0.95
```

Note that, while it is often claimed that Bayesian 95% credible intervals that do not cross zero can be used to indicate statistical significance in the classical (Frequentist) sense, this does not hold for variance components here as they are constrained to be positive in MCMCglmm. As such, a lower bound of the credible interval close to zero might actually indicate low confidence in a non-zero proportion of the phenotypic variance in exploration being explained by differences among individuals.

Let's do the same for boldness:

```
mcmc_prop_B <- mcmc_E_B_us$VCV[,"traitboldness:traitboldness.ID"]/(
  mcmc_E_B_us$VCV[,"traitboldness:traitboldness.ID"] +
    mcmc_E_B_us$VCV[,"traitboldness:traitboldness.units"]
)

mean(mcmc_prop_B)
```

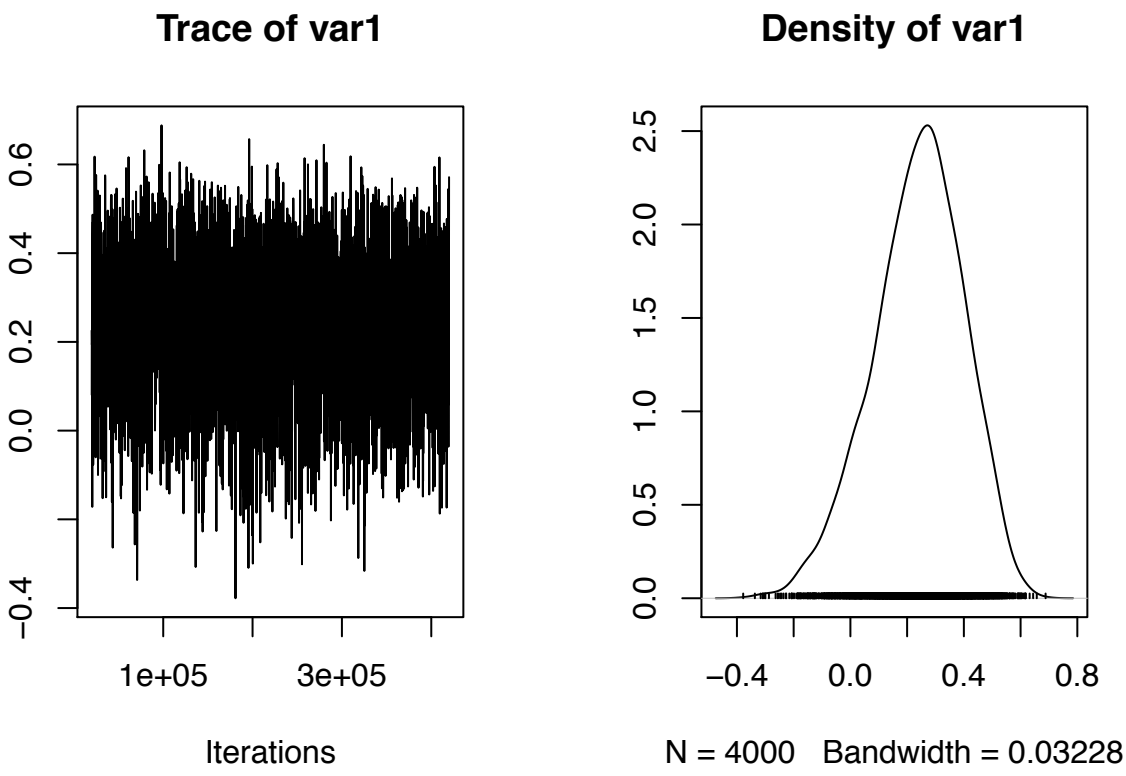
```
## [1] 0.3751389
```

```
HPDinterval(mcmc_prop_B)
```

```
##          lower     upper
## var1 0.2602269 0.4977966
## attr(,"Probability")
## [1] 0.95
```

We can also use this process to estimate the mean and credible intervals of the correlation from our model (co)variances. We create a posterior distribution of the among-individual correlation by dividing the corresponding covariance between boldness and exploration by the product of the square root of their variances (i.e., standardising the covariance to a scale from -1 to 1):

```
mcmc_cor_EB <- mcmc_E_B_us$VCV[,"traitboldness:traitexploration.ID"]/
  (sqrt(mcmc_E_B_us$VCV[,"traitboldness:traitboldness.ID"])*
     sqrt(mcmc_E_B_us$VCV[,"traitexploration:traitexploration.ID"]))

plot(mcmc_cor_EB)
```




```
mean(mcmc_cor_EB)
```

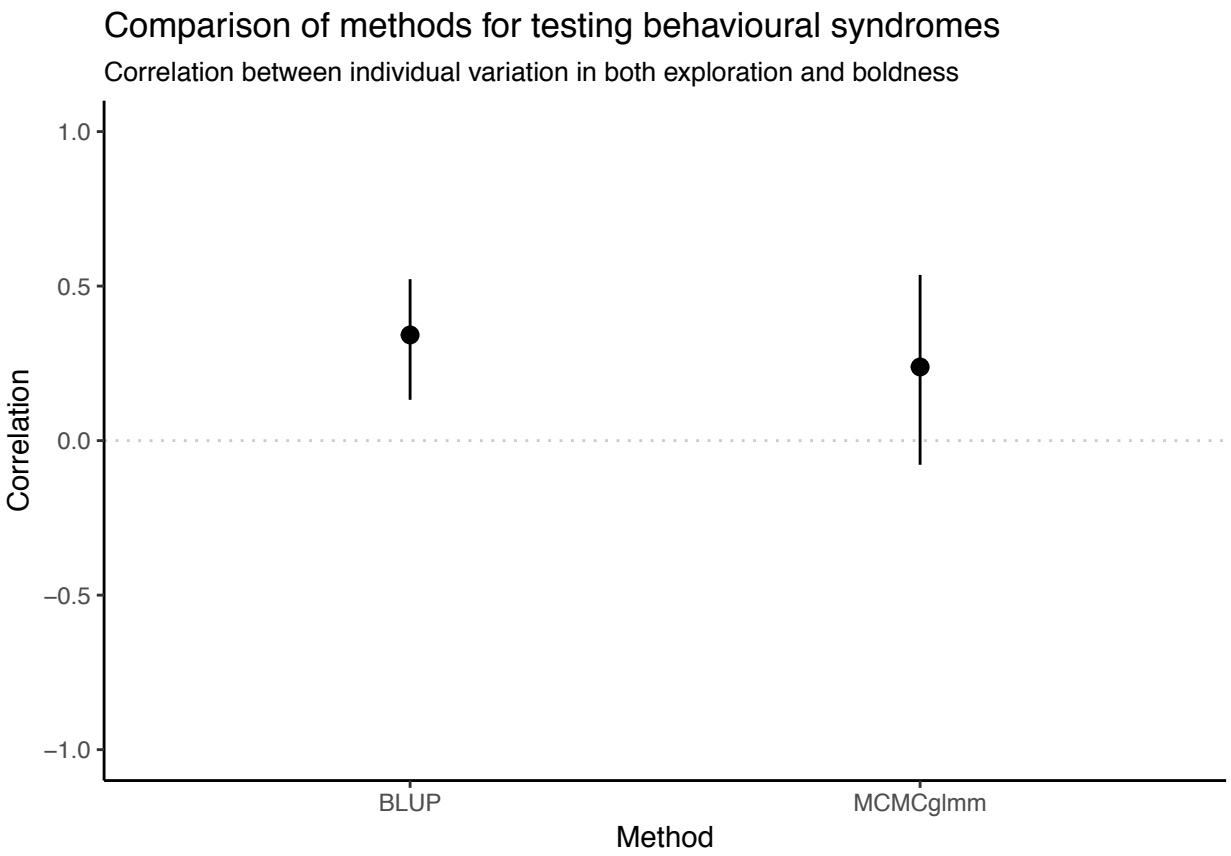
```
## [1] 0.2383352
```

```
HPDinterval(mcmc_cor_EB)
```

```
##            lower    upper
## var1 -0.07829537 0.536206
## attr(,"Probability")
## [1] 0.95
```

In this case, because the correlation can take on either positive or negative then we can use the credible interval to assess statistical significance. Here the 95% credible interval spans zero, and since the model fit is good, we should conclude that there is no evidence of a statistically significant correlation.

To better demonstrate that BLUPs produce anticonservative hypothesis tests, we can plot the correlation estimates and their confidence/credible intervals from the two approaches that we have taken. The CI are taken directly from the `cor.test` function for the BLUPs, and for `MCMCglmm` they are taken from the posterior distribution of correlation samples (using the `HPDinterval` function).



Here we can clearly see that the BLUPs method - having failed to carry through the error around the predictions of individual-level estimates - is anticonservative, with small confidence intervals (and a correspondingly small P-value, $P = 0.0019$). Testing the syndrome directly in a bivariate model that retains all the data, by comparison, enables us to capture the true uncertainty about the estimate of the correlation. This is reflected in the larger CI which, in this case, cross zero and thus indicate a lack of support for a statistically significant behavioural syndrome.

Adding further traits

As part of our data collection, we also have a single value of mating success for each individual (which we will use as a proxy for fitness). We are interested in how our personality traits correlate with variation in this fitness-related measure. While our test above showed that the correlation between the measured personality traits was not significant, there did appear to be some relationship — so we shall incorporate both personality traits and fitness into a single trivariate model for hypothesis testing.

In this case, because the new response variable to be added to our model is fitness, we are **not** going to mean-centre and scale by phenotypic standard deviations, but instead divide by the mean fitness value (such that we are investigating among-individual covariance between personality traits and **relative fitness**). We create this new variable, `rel_fitness`, as follows:

```
df_syndrome <- df_syndrome %>%
  mutate(rel_fitness = fitness/mean(fitness, na.rm=TRUE))
```

Note that we will refer to this relative fitness trait simply as 'fitness' below for simplicity's sake.

Setting up the model

Below, we will set up our main model, which will allow for heterogeneous among-individual variances in our 3 traits (boldness, exploration, fitness), and will estimate the covariance between them.

First, we set up a prior, which we specify in a similar way as the bivariate model. However, for the residuals (or 'within-individual' variance), we must bear in mind that we have only a single fitness value per individual — therefore, that trait has **no within-individual variance**, and **within-individual correlations involving fitness must be 0**. We can set the variance component to a particular value using the `fix` command, although as variances have to be positive we fix the within-individual variance in fitness to a small positive number (here, 0.0001):

```
prior_E_B_fit_1px = list(R = list(V = diag(c(1,1,0.0001),3,3), nu = 1.002, fix = 3),
                         G = list(G1 = list(V = diag(3), nu = 3,
                                            alpha.mu = rep(0,3),
                                            alpha.V = diag(25^2,3,3))))
```

Now, we can fit our model with these starting values and constraints. Again, we `cbind` our response variables on the left-hand side of the formula, and use `trait` to denote a multivariate model. We can also use the `at.level` keyword to specify that fixed effects are estimated only for certain traits — here, we test for an effect of assay repeat only on exploration and boldness (because these were measured repeatedly), while we test for the effect of body size on all of our traits.

Note that in the model specification below, we set the argument `pr = TRUE`*. This saves the posterior distribution of the individual random effects (analagous to the BLUP from the REML analysis) so we can visualise them later, but does take up more memory (over 8Mb compared to <1Mb for a model run without saving these values).*

Fit the model as follows (and be sure to use diagnostic checks). Note that I have increased the number of iterations (and both the burnin and thinning interval), so once it's underway, that's a good time to go and make a cup of tea... (the run will likely take over 20 minutes).

```
mcmc_E_B_fit <- MCMCglmm(cbind(scale(exploration),
                               scale(boldness),
                               rel_fitness) ~ trait-1 +
                           at.level(trait,1):scale(assay_rep, scale = FALSE) +
                           at.level(trait,2):scale(assay_rep, scale = FALSE) +
                           trait:scale(body_size),
                         random =~  us(trait):ID,
                         rcov =~ us(trait):units,
                         family = c("gaussian","gaussian","gaussian"),
                         prior = prior_E_B_fit_1px,
                         nitt=750000,
                         burnin=50000,
                         thin=175,
                         verbose = TRUE,
                         pr = TRUE,
                         data = as.data.frame(df_syndrome))
```

Take a look at the model summary:

```
summary(mcmc_E_B_fit)
```

As before, we get (co)variance estimates, credible intervals, and effective sample sizes for the among-individual and residual variance terms. Note that our constraint on the residual ('within-individual') variance term for our fitness measure: the `rel_fitness:rel_fitness.units` estimate is at 0.0001, with an effective sample

size of 0. You should also note that the within-individual covariance terms involving the fitness trait are very close to 0, with very small effective sample sizes, so the model has effectively not fit these covariances (which is what we wanted).

A quick sanity check also tells us that the correlation between boldness and exploration estimated in this model is the same as in our earlier bivariate model:

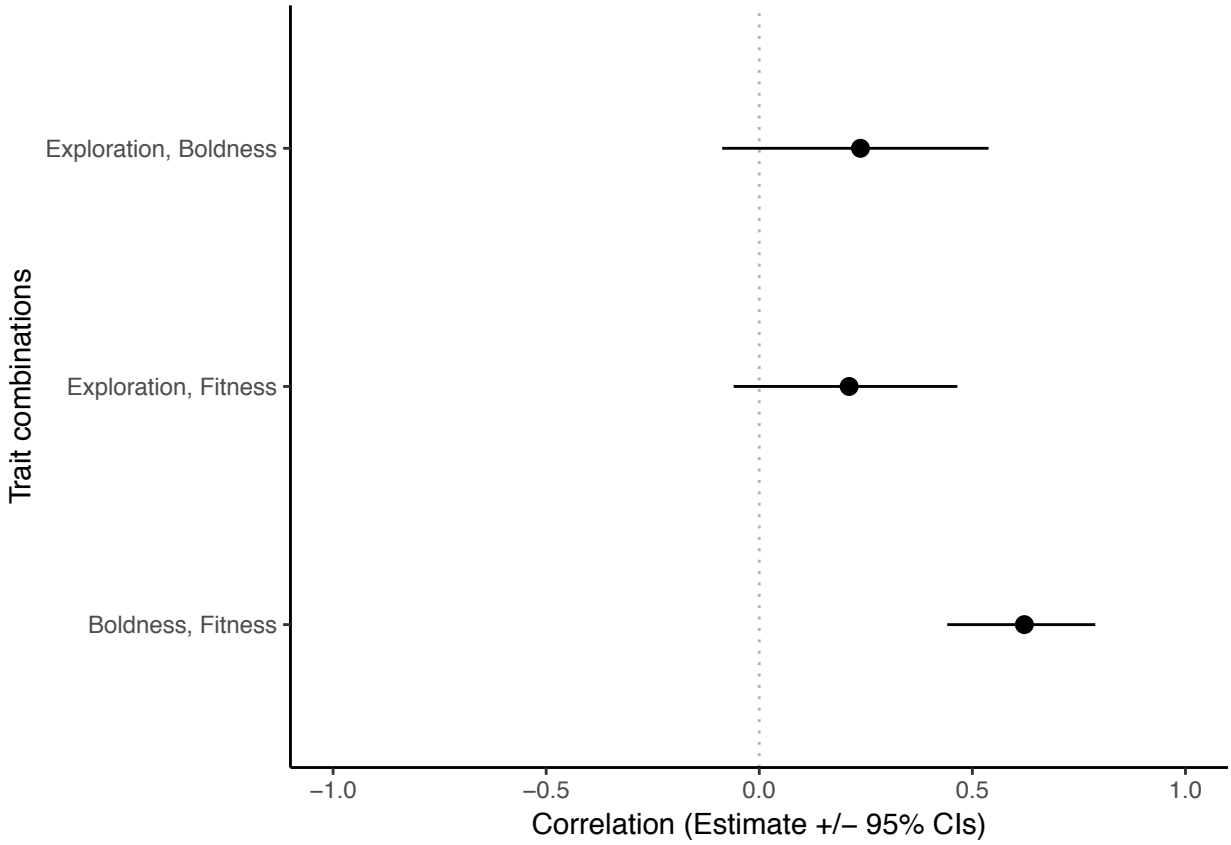
```
mcmc_E_B_fit_cor_EB <- mcmc_E_B_fit$VCV[,"traitboldness:traitexploration.ID"]/
  (sqrt(mcmc_E_B_fit$VCV[,"traitboldness:traitboldness.ID"])*
     sqrt(mcmc_E_B_fit$VCV[,"traitexploration:traitexploration.ID"]))

mean(mcmc_E_B_fit_cor_EB)
HPDinterval(mcmc_E_B_fit_cor_EB)
```

```
## [1] 0.2374761
##            lower     upper
## var1 -0.08700906 0.5379599
## attr(,"Probability")
## [1] 0.95
```

As before, we can use our posterior distributions to estimate the among-individual correlations between each of our traits of interest, and assess statistical significance using their 95% credible intervals from our `MCMCglmm` model:

```
mcmc_E_B_fit_cor_EB <- mcmc_E_B_fit$VCV[,"traitboldness:traitexploration.ID"]/
  (sqrt(mcmc_E_B_fit$VCV[,"traitboldness:traitboldness.ID"])*
     sqrt(mcmc_E_B_fit$VCV[,"traitexploration:traitexploration.ID"]))

mcmc_E_B_fit_cor_Efit <- mcmc_E_B_fit$VCV[,"traitrel_fitness:traitexploration.ID"]/
  (sqrt(mcmc_E_B_fit$VCV[,"traitrel_fitness:traitrel_fitness.ID"])*
     sqrt(mcmc_E_B_fit$VCV[,"traitexploration:traitexploration.ID"]))

mcmc_E_B_fit_cor_Bfit <- mcmc_E_B_fit$VCV[,"traitrel_fitness:traitboldness.ID"]/
  (sqrt(mcmc_E_B_fit$VCV[,"traitrel_fitness:traitrel_fitness.ID"])*
     sqrt(mcmc_E_B_fit$VCV[,"traitboldness:traitboldness.ID"]))

df_mcmc_cors <- data_frame(Traits = c("Exploration, Boldness",
                                      "Exploration, Fitness",
                                      "Boldness, Fitness"),
                           Estimate = c(mean(mcmc_E_B_fit_cor_EB),
                                        mean(mcmc_E_B_fit_cor_Efit),
                                        mean(mcmc_E_B_fit_cor_Bfit)),
                           Lower = c(HPDinterval(mcmc_E_B_fit_cor_EB)[,"lower"],
                                     HPDinterval(mcmc_E_B_fit_cor_Efit)[,"lower"],
                                     HPDinterval(mcmc_E_B_fit_cor_Bfit)[,"lower"]),
                           Upper = c(HPDinterval(mcmc_E_B_fit_cor_EB)[,"upper"],
                                     HPDinterval(mcmc_E_B_fit_cor_Efit)[,"upper"],
                                     HPDinterval(mcmc_E_B_fit_cor_Bfit)[,"upper"]))

ggplot(df_mcmc_cors, aes(x = Traits, y = Estimate)) +
  geom_pointrange(aes(ymin = Lower,
                      ymax = Upper)) +
  geom_hline(yintercept = 0,
             linetype = "dotted",
```

```
            alpha = 0.3) +
scale_x_discrete(limits = c("Boldness, Fitness",
                            "Exploration, Fitness",
                            "Exploration, Boldness")) +
labs(x = "Trait combinations",
     y = "Correlation (Estimate +/- 95% CIs)") +
ylim(-1,1) +
coord_flip() +
theme_classic()
```



We might want to present our final results as a matrix with variances on the diagonals, covariances below and correlations above (with the lower and upper bounds of 95% CIs in parentheses):

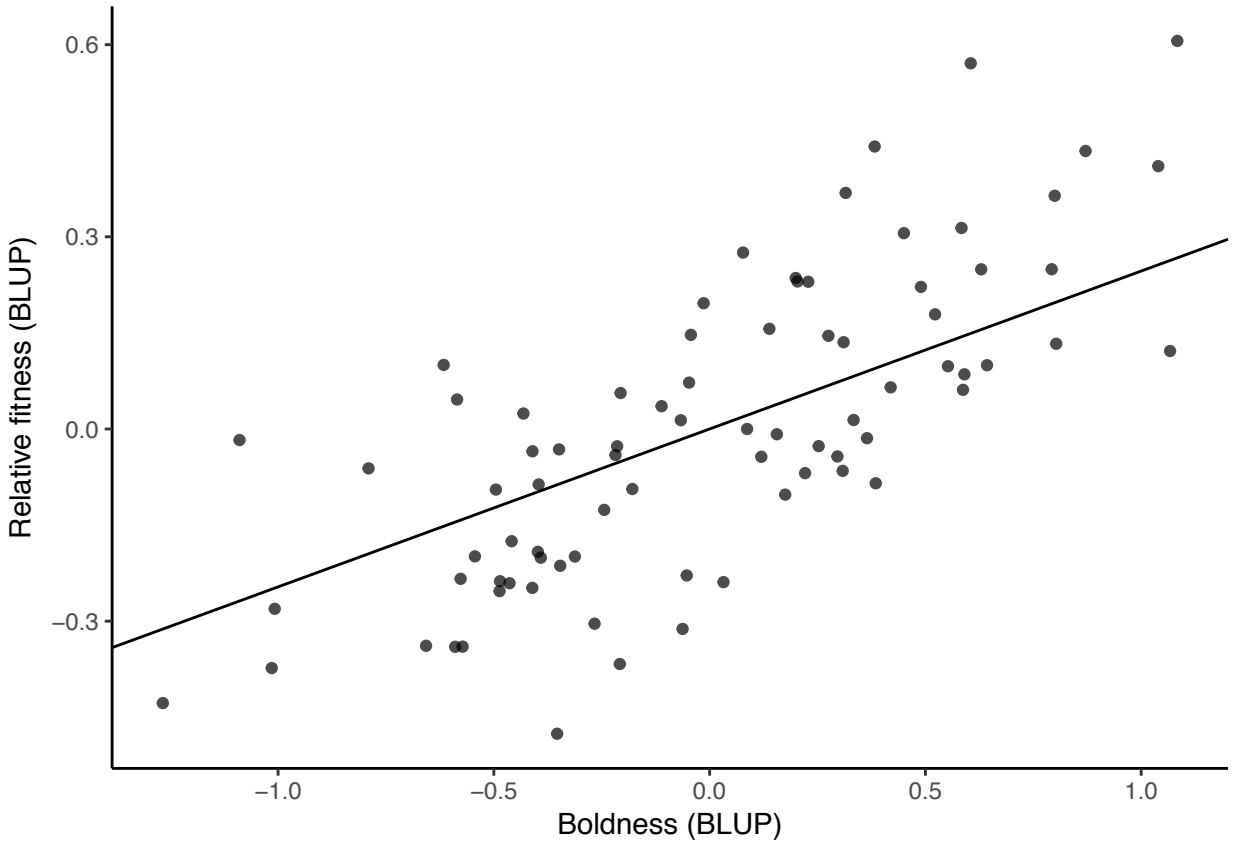
	Exploration	Boldness	Fitness
Exploration	0.29 (0.13,0.45)	0.24 (-0.09,0.54)	0.21 (-0.06,0.46)
Boldness	0.08 (-0.04,0.21)	0.39 (0.22,0.57)	0.62 (0.44,0.79)
Fitness	0.03 (-0.01,0.07)	0.09 (0.05,0.14)	0.06 (0.04,0.08)

Conclusions

To conclude, then: we found that the correlation between boldness and exploration tends to be positive among male haggis, but this correlation is not statistically significant and thus does not provide strong evidence for a behavioural syndrome. However, inappropriate analysis of BLUP extracted from univariate models would lead to a different (erroneous) conclusion. We also found no statistically significant association between among-individual variation in exploration and fitness. However, we did find a statistically significant positive association between among-individual variation in boldness and our fitness proxy, indicating that bolder male haggis had greater mating success (see figure below).

Note: below, we use posterior modes of random effects (BLUPs from the MCMCglmm model) from our trivariate model to construct a figure that illustrates the association between boldness and fitness. Unlike its use in secondary statistical analyses, this is an appropriate use of BLUPs — i.e., just for illustrative purposes!

```
df_bf_coefs <- data_frame(Trait = attr(colMeans(mcmc_E_B_fit$Sol), "names"),
                          Value = colMeans(mcmc_E_B_fit$Sol)) %>%
  separate(Trait, c("Trait","Type","ID"), sep = "\\.", fill = "right") %>%
  filter(Type == "ID") %>%
  filter(Trait %in% c("traitboldness", "traitrel_fitness")) %>%
  select(-Type) %>%
  spread(Trait, Value)


# Find the regression line -
#  the covariance of boldness, relative fitness divided by
#  the square root of the variance in boldness
B_fit_slope <-  mcmc_E_B_fit$VCV[,"traitrel_fitness:traitboldness.ID"]/
  mcmc_E_B_fit$VCV[,"traitboldness:traitboldness.ID"]


ggplot(df_bf_coefs, aes(x = traitboldness, y = traitrel_fitness, group = ID)) +
  geom_point(alpha = 0.7) +
  geom_abline(intercept = 0, slope = mean(B_fit_slope)) +
  labs(x = "Boldness (BLUP)",
       y = "Relative fitness (BLUP)") +
  theme_classic()
```



Further tutorials

We will continue to develop tutorials for multivariate modelling of individual (co)variation, which will cover some of the more advanced issues discussed in our paper. Please visit tomhouslay.com for more information.

ID	assay_rep	boldness	exploration	fitness		body_size
S_1	1	18.5745096	39.7364776		39	21.7179479
S_1	2	18.3187658	39.4075446	NA		21.545649
S_1	3	20.3325558	40.1580349	NA		21.3416353
S_1	4	19.4049967	40.2904165	NA		20.7776099
S_2	1	20.6978577	39.4682008		56	25.7079314
S_2	2	18.5963845	40.119347	NA		26.4266061
S_2	3	22.2375321	41.3469193	NA		26.0968418
S_2	4	19.813581	41.2340122	NA		25.746078
S_3	1	20.3500735	38.5867528		51	29.1519091
S_3	2	19.6911677	40.9543552	NA		28.9814442
S_3	3	21.1541311	41.2069454	NA		29.1014692
S_3	4	20.0333852	39.8195935	NA		29.0321869
S_4	1	17.8933417	39.4149024		31	26.2441252
S_4	2	19.5140168	40.5488653	NA		25.5521613
S_4	3	18.8620161	42.1361511	NA		25.950951
S_4	4	21.9650019	41.6851158	NA		25.3195146
S_5	1	19.3589778	37.3041546		39	25.1542209
S_5	2	21.8242363	38.9038413	NA		24.8381485
S_5	3	18.4131583	36.2623051	NA		24.394314
S_5	4	20.5859026	39.7551139	NA		24.8722273
S_6	1	18.009306	40.7869754		30	16.4566359
S_6	2	20.141488	40.4712413	NA		17.5762429
S_6	3	17.7786	38.6306008	NA		16.6604519
S_6	4	18.9872707	38.7975312	NA		16.8105215
S_7	1	20.4255094	40.8399185		47	23.5527359
S_7	2	19.5870014	41.1842695	NA		23.2531662
S_7	3	21.7336751	39.9017845	NA		23.8345554
S_7	4	18.7230737	39.3023959	NA		23.6587904
S_8	1	18.1343989	40.6229724		21	23.5037662
S_8	2	20.2412012	39.7360733	NA		23.5902292
S_8	3	22.3061942	42.332719	NA		23.3977746
S_8	4	19.0649385	38.299801	NA		22.7857401
S_9	1	22.5846511	41.016438		45	25.4145859
S_9	2	19.7797451	40.6627197	NA		25.0603701
S_9	3	20.0940153	39.3067831	NA		25.0353419
S_9	4	22.6817987	41.8222173	NA		26.6601099
S_10	1	18.4984294	38.9234546		27	26.0698439
S_10	2	18.6855911	39.6788989	NA		26.8457464
S_10	3	19.5780368	40.056343	NA		25.9171968
S_10	4	20.2166465	40.0496865	NA		26.4060506
S_11	1	22.3203978	40.87835		64	22.9987653
S_11	2	20.488988	39.1084831	NA		23.9712535
S_11	3	20.2272368	40.3719119	NA		23.704132
S_11	4	19.051995	39.8994974	NA		23.9304685
S_12	1	20.916161	37.8143928		52	32.7614238

S_12	2	21.2770301	40.4680074	NA		32.4773686
S_12	3	22.2734238	38.8686762	NA		33.0374506
S_12	4	20.4107842	38.5833465	NA		32.9201407
S_13	1	20.5386633	39.7121528		50	23.0771224
S_13	2	20.3235903	39.4927472	NA		23.6781737
S_13	3	20.7902295	40.2186811	NA		24.2548167
S_13	4	18.7961807	40.5307496	NA		23.8086135
S_14	1	21.3780354	40.7691199		47	26.4406121
S_14	2	21.2708812	40.6684207	NA		26.070472
S_14	3	20.246382	40.1886095	NA		26.0814769
S_14	4	19.5713066	41.2991931	NA		25.4902434
S_15	1	19.1361122	40.7216586		41	22.9172444
S_15	2	21.4128396	40.953385	NA		22.2951542
S_15	3	19.5018127	39.5367991	NA		23.5229609
S_15	4	19.4689557	40.7286903	NA		23.0104524
S_16	1	18.3709747	38.5975406		39	22.6159207
S_16	2	18.2602969	40.0277989	NA		23.0747574
S_16	3	19.5273619	40.5583875	NA		23.8301403
S_16	4	20.0514375	39.3244004	NA		23.0498418
S_17	1	18.4468357	39.7205978		42	27.2203793
S_17	2	20.6879928	40.5566464	NA		26.6533751
S_17	3	18.9820085	39.5616643	NA		28.4670179
S_17	4	18.8036958	42.445244	NA		27.5201417
S_18	1	18.5084248	39.9763869		27	25.6497336
S_18	2	19.1832264	40.4481508	NA		25.8910715
S_18	3	20.6914284	40.1878126	NA		25.3810824
S_18	4	17.9961485	39.5515537	NA		25.0694127
S_19	1	22.3440192	41.1070143		45	25.0833317
S_19	2	22.8488818	42.7382517	NA		25.7779637
S_19	3	20.144399	42.8041754	NA		25.3184728
S_19	4	21.5273713	41.7605539	NA		24.9896374
S_20	1	18.8523067	41.8752684		38	23.1475075
S_20	2	17.1208331	38.7618034	NA		23.0809379
S_20	3	17.375703	38.5178718	NA		23.8355127
S_20	4	19.7904725	40.5153615	NA		23.6152005
S_21	1	21.6086147	42.7233141		59	28.9337998
S_21	2	22.3003042	41.298978	NA		29.5539935
S_21	3	21.810981	40.6775149	NA		28.6157213
S_21	4	21.1964817	40.5888089	NA		28.5959927
S_22	1	19.8254843	39.1494329		46	22.81607
S_22	2	20.5496198	39.479239	NA		22.6040782
S_22	3	21.0938151	39.8256834	NA		21.8346189
S_22	4	19.9120542	39.1853006	NA		21.8228173
S_23	1	18.0149032	39.2284674		40	24.3581383
S_23	2	19.2293286	41.337362	NA		24.1526271
S_23	3	15.7439067	39.3420791	NA		24.3806926

S_23	4	17.3814485	40.3025255	NA		24.3218634
S_24	1	21.4973584	41.5354146		45	19.2161799
S_24	2	21.6099297	42.4086221	NA		18.5437906
S_24	3	22.9377278	40.8084415	NA		19.7044647
S_24	4	23.368499	42.4713714	NA		19.4166925
S_25	1	21.8619657	40.0179747		37	21.6455196
S_25	2	21.0752545	39.8229201	NA		21.3603348
S_25	3	22.0616256	39.8915512	NA		21.4857507
S_25	4	18.1342232	37.3411701	NA		20.6437658
S_26	1	19.8576008	39.3372583		46	18.6254355
S_26	2	18.1407081	39.4438723	NA		19.4685917
S_26	3	20.4195346	40.595597	NA		19.2455343
S_26	4	19.4598984	39.8434782	NA		19.1943318
S_27	1	19.502591	37.7527511		36	20.4070011
S_27	2	20.5732658	40.833898	NA		21.2830583
S_27	3	20.1493681	39.4876244	NA		19.5125502
S_27	4	21.2462382	40.8863187	NA		20.2980269
S_28	1	20.2277329	41.3580191		38	25.4408008
S_28	2	21.9951871	42.927774	NA		25.9762737
S_28	3	20.9429167	41.1652494	NA		24.3834493
S_28	4	20.7032575	41.9741604	NA		24.0029807
S_29	1	21.2588113	40.6823763		59	25.9407319
S_29	2	20.9103901	40.6907624	NA		26.9986345
S_29	3	19.0080703	39.9047801	NA		26.1231398
S_29	4	20.1445535	39.9512805	NA		26.5307893
S_30	1	20.692202	40.9099058		50	26.2061063
S_30	2	20.5149804	40.2160333	NA		26.7553967
S_30	3	22.3364904	42.4616241	NA		26.0819035
S_30	4	20.6977448	41.2005983	NA		26.1456003
S_31	1	21.3201958	42.6417435		42	22.2939361
S_31	2	19.07113	38.9492296	NA		22.8380021
S_31	3	19.1160588	41.4728138	NA		22.8345286
S_31	4	19.6620312	40.6061371	NA		22.8585925
S_32	1	19.6148146	41.8326281		40	27.7578283
S_32	2	19.7694582	40.6883783	NA		28.0838238
S_32	3	20.6683368	41.901706	NA		26.3517306
S_32	4	19.6858676	40.9737694	NA		27.9911355
S_33	1	20.5394508	38.7238817		36	26.9765309
S_33	2	19.0660448	40.5251259	NA		26.7234899
S_33	3	19.2981651	40.5664434	NA		27.3390189
S_33	4	20.052361	39.7043521	NA		27.7561844
S_34	1	18.5857254	39.9425704		34	28.5278363
S_34	2	20.6714988	39.7565401	NA		28.9071031
S_34	3	18.554187	38.3266427	NA		29.1833293
S_34	4	19.4039804	38.8015043	NA		29.2625579
S_35	1	20.6134521	40.4990448		37	24.6477249

S_35	2	18.1453723	39.581873	NA		25.9221668
S_35	3	19.2172773	38.8116809	NA		24.8213931
S_35	4	18.25196	40.7543482	NA		25.5496639
S_36	1	22.0597752	40.1295647		66	28.0596096
S_36	2	22.0531556	42.2040191	NA		28.296935
S_36	3	21.8314653	41.7000091	NA		28.8406093
S_36	4	21.6286812	40.0086683	NA		28.5884584
S_37	1	19.1497155	41.8134205		33	28.3239569
S_37	2	20.1869353	40.8218882	NA		27.1004888
S_37	3	20.2587218	40.8430038	NA		27.4458581
S_37	4	19.9620661	39.7979772	NA		27.9175578
S_38	1	19.9445386	40.9054105		31	28.9856891
S_38	2	20.8406704	39.7555277	NA		29.0507795
S_38	3	19.0029315	40.6271079	NA		29.0901373
S_38	4	18.601368	41.724883	NA		28.9148778
S_39	1	19.2995925	41.1528384		37	25.8457506
S_39	2	20.0769231	39.7617934	NA		26.0652528
S_39	3	19.5729128	39.0193976	NA		25.6507338
S_39	4	20.393454	40.2940586	NA		26.679179
S_40	1	21.2288126	39.8000424		53	22.3582569
S_40	2	21.389392	41.743212	NA		21.3833909
S_40	3	20.105638	39.7672643	NA		21.7924235
S_40	4	19.1392192	38.0115736	NA		21.7735803
S_41	1	18.4680403	40.4288298		26	28.6146864
S_41	2	19.5561695	42.098974	NA		28.3118565
S_41	3	17.8213072	40.0901952	NA		28.5193823
S_41	4	18.6002561	41.0849123	NA		29.0736105
S_42	1	21.2928159	39.6329764		51	30.6180627
S_42	2	19.8445145	38.8588315	NA		31.0314399
S_42	3	19.6575758	39.9218787	NA		31.0571501
S_42	4	20.5499385	39.6059796	NA		30.5862105
S_43	1	20.1849537	41.0842111		31	24.9851652
S_43	2	21.2768262	39.6057029	NA		24.5869981
S_43	3	18.3021585	39.35726	NA		23.4291114
S_43	4	22.3321302	41.2459695	NA		24.9110991
S_44	1	21.7624436	41.3757242		41	21.4052791
S_44	2	20.6590816	40.4953954	NA		22.5801939
S_44	3	20.2180356	40.2963319	NA		22.4817919
S_44	4	20.539635	41.1723482	NA		21.1260063
S_45	1	22.2975669	39.937893		40	25.9808267
S_45	2	19.9698291	40.2629308	NA		26.3331991
S_45	3	20.2612259	38.5820382	NA		25.4255904
S_45	4	20.0188436	39.922297	NA		26.151356
S_46	1	19.0144609	39.5142077		25	19.368596
S_46	2	20.8073106	39.5590644	NA		20.6594759
S_46	3	18.4199352	38.4074998	NA		20.8986507

S_46	4	21.2760893	40.8344627	NA		20.2981395
S_47	1	20.5887775	41.4691403		29	29.4788895
S_47	2	20.4023763	39.7108087	NA		30.4887379
S_47	3	21.1615313	40.690556	NA		30.5878332
S_47	4	18.7014666	40.6759218	NA		29.8278456
S_48	1	19.4978762	39.6621512		55	20.8734015
S_48	2	22.3547563	40.6364569	NA		21.5598907
S_48	3	22.094183	40.2100389	NA		21.6397673
S_48	4	20.7556076	40.0128673	NA		20.7783872
S_49	1	20.3774009	39.8408843		52	24.3653619
S_49	2	20.6808871	39.4809421	NA		24.9358099
S_49	3	18.8550945	39.8257659	NA		24.4806621
S_49	4	18.8543332	38.4074488	NA		24.6790708
S_50	1	20.0007744	40.290888		28	25.0146284
S_50	2	20.5558534	38.6495866	NA		24.7835962
S_50	3	20.2308128	37.8708165	NA		25.8727329
S_50	4	19.8447746	37.7694342	NA		25.4037196
S_51	1	19.0968304	39.4389557		32	24.7340385
S_51	2	19.6018119	40.7184445	NA		24.6067265
S_51	3	20.4365402	41.416918	NA		24.8818508
S_51	4	19.326485	39.3206843	NA		24.2907087
S_52	1	19.9112334	39.509466		33	28.0003697
S_52	2	19.0696147	39.4836549	NA		28.0660459
S_52	3	18.7247388	38.778721	NA		27.796475
S_52	4	20.1770585	39.2793432	NA		28.1789759
S_53	1	22.1378293	40.6655696		40	22.3024905
S_53	2	18.7001698	39.7760533	NA		22.5700932
S_53	3	20.4843735	40.2926382	NA		23.7575723
S_53	4	22.1023138	40.5374088	NA		23.3147724
S_54	1	23.5815088	40.943509		58	27.1070541
S_54	2	22.2629885	40.6548263	NA		27.3655635
S_54	3	19.6578182	39.1890904	NA		27.0871502
S_54	4	22.1682406	40.0181649	NA		27.6900935
S_55	1	19.7143376	39.8461339		48	24.9376305
S_55	2	22.649033	41.9762808	NA		25.338713
S_55	3	22.3919471	41.6142071	NA		25.1442628
S_55	4	19.5405895	39.3764171	NA		24.4667415
S_56	1	20.5903599	41.4057226		51	25.8672306
S_56	2	20.5073512	40.7021683	NA		26.4553518
S_56	3	22.3001959	40.0457158	NA		26.6547387
S_56	4	22.9289077	39.5607187	NA		25.3058832
S_57	1	19.176239	41.6951245		43	25.0607041
S_57	2	19.0491262	39.1572519	NA		23.9508961
S_57	3	19.3824122	38.4743878	NA		25.0714503
S_57	4	20.1619877	40.3575801	NA		23.6431026
S_58	1	19.4236559	40.3886091		34	33.3693868

S_58	2	19.7550135	39.8554965	NA		32.6400291
S_58	3	20.2334749	39.7166257	NA		33.7655239
S_58	4	19.5800778	40.2328857	NA		33.9418998
S_59	1	18.9295295	39.1930496		44	27.4522471
S_59	2	20.4581372	41.2351348	NA		26.540432
S_59	3	19.4193763	40.9943312	NA		27.4488022
S_59	4	21.2251865	40.705727	NA		28.5683628
S_60	1	20.5357082	40.8832625		39	31.908719
S_60	2	22.758753	42.3737861	NA		31.9644417
S_60	3	21.7423228	41.6430006	NA		32.4992143
S_60	4	20.6847399	42.2766288	NA		33.0940735
S_61	1	19.2535351	40.8172985		42	19.6593794
S_61	2	17.7685247	39.2853262	NA		19.0660723
S_61	3	19.0271342	41.4456123	NA		20.0168603
S_61	4	18.1371317	40.6025447	NA		20.1768535
S_62	1	20.1951871	40.9454921		41	26.7213667
S_62	2	19.6623252	40.1097315	NA		26.240884
S_62	3	21.0342082	40.7322084	NA		26.3164773
S_62	4	20.8625273	41.7290532	NA		25.334772
S_63	1	20.6370998	39.1966131		44	22.9201001
S_63	2	23.0424545	41.3044778	NA		22.9358703
S_63	3	21.4975511	39.5803081	NA		22.9615912
S_63	4	19.3173238	38.9680609	NA		22.432261
S_64	1	18.4306837	38.3369116		31	24.8176805
S_64	2	19.3484425	39.4375111	NA		25.1126015
S_64	3	18.9362889	39.523328	NA		25.6210151
S_64	4	20.0816698	40.1515142	NA		24.9801287
S_65	1	19.1674846	39.6048063		37	23.1529361
S_65	2	19.2231127	39.8527594	NA		23.0426631
S_65	3	18.8573102	40.5452862	NA		23.4275373
S_65	4	19.7057919	40.4705204	NA		22.890896
S_66	1	19.2463595	39.8547051		31	27.6783358
S_66	2	19.6394158	41.4156703	NA		27.2219161
S_66	3	20.6777186	42.202508	NA		27.4502497
S_66	4	19.6907302	42.2489029	NA		27.2374271
S_67	1	19.0560628	39.1109065		39	24.1717438
S_67	2	19.8639552	39.6392123	NA		23.2380469
S_67	3	21.5683817	42.3618508	NA		23.4784784
S_67	4	17.7998906	38.9961484	NA		23.0785917
S_68	1	21.1596608	40.7997976		54	27.7603871
S_68	2	23.8192821	43.7418223	NA		26.5709924
S_68	3	20.5485087	40.2938102	NA		27.520562
S_68	4	19.3367543	40.4458126	NA		26.1050496
S_69	1	16.3175646	40.1138905		23	23.8576247
S_69	2	16.9366895	39.4797416	NA		23.9787928
S_69	3	19.7123602	42.0170119	NA		24.4422892

S_69	4	18.6691016	41.5634439	NA		23.9314276
S_70	1	20.1774683	41.7842506		43	22.0684405
S_70	2	22.5010197	40.5540767	NA		22.8907045
S_70	3	21.4573296	40.2377738	NA		21.697342
S_70	4	19.4907664	40.0390903	NA		23.1631007
S_71	1	22.062817	42.0174631		49	25.1943235
S_71	2	17.7658367	39.7318657	NA		25.0526643
S_71	3	19.7071026	40.193499	NA		25.5161551
S_71	4	19.5320223	39.184341	NA		24.8793494
S_72	1	19.9100242	40.2159084		45	25.0998485
S_72	2	18.5740167	41.994849	NA		25.0227133
S_72	3	17.5496783	40.3524862	NA		25.8165705
S_72	4	18.6021702	40.9439158	NA		25.6335907
S_73	1	18.7177056	39.9525557		33	28.059577
S_73	2	18.5801954	38.559737	NA		27.658666
S_73	3	18.9881804	39.5305029	NA		28.259371
S_73	4	20.2176806	38.9252541	NA		27.1388042
S_74	1	20.3604129	38.672499		39	24.6850117
S_74	2	20.5690814	39.5594288	NA		24.946812
S_74	3	20.8127676	39.2293718	NA		23.1206498
S_74	4	20.5498527	38.5193001	NA		23.4497805
S_75	1	21.8532654	39.9351787		41	28.0999169
S_75	2	19.662919	38.8254437	NA		28.7459077
S_75	3	19.983385	39.4359793	NA		28.9540206
S_75	4	20.4493322	40.2944109	NA		28.9164229
S_76	1	21.5266738	41.0965849		43	22.3074762
S_76	2	21.1297881	39.9729223	NA		22.6326136
S_76	3	21.9489244	40.4288028	NA		21.7283838
S_76	4	20.6470031	41.5978915	NA		23.0183813
S_77	1	21.5147566	41.689573		46	23.2624455
S_77	2	23.4354305	41.4365826	NA		23.7641727
S_77	3	20.7856487	39.9105626	NA		23.1857809
S_77	4	21.2144695	40.7338936	NA		23.0510163
S_78	1	18.3567337	38.2968853		29	22.3338277
S_78	2	18.24404	39.0296846	NA		21.3077351
S_78	3	17.4822929	37.9992596	NA		21.7100891
S_78	4	17.4453001	40.0366282	NA		20.9343266
S_79	1	21.025615	41.0589345		32	28.8089343
S_79	2	20.1115862	40.2807299	NA		29.9064397
S_79	3	20.4226393	39.1261719	NA		28.7722819
S_79	4	19.9817992	38.6207354	NA		29.2348008
S_80	1	19.2141392	40.041069		27	25.7609549
S_80	2	18.8618925	39.3839958	NA		26.4205213
S_80	3	19.226019	41.0399314	NA		25.7012092
S_80	4	19.8660756	39.9686084	NA		26.5435477