

**Bayesian Analysis of Heterogeneous Residual Variance in Canine Behaviour**

**S.M. van den Berg,\* I. Schwabe,\* W.F. Fikse,† H.C.M. Heuven,‡ C.A.W. Glas\***

\*Behavioural Sciences, University of Twente, Enschede, The Netherlands, †Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden, ‡Veterinary Sciences, University of Utrecht, Utrecht, The Netherlands

**ABSTRACT:** Non-additive effects like for instance genotype-environment interactions and genetically structured heterogeneity of residual variance are notoriously dependent on scale: many statistical non-additive phenomena disappear after a careful choice of transformation of the phenotypic values. Particularly for behavioural measures, scale is a delicate matter. We present a novel Bayesian approach that assesses heterogeneity in environmental variance as a function of genetic effects, where the scale is defined by a psychometric model based on item-response theory. This makes analysis results independent of what items are in a particular test version. We apply the method to fearful behaviour in dogs and compare results with the more usual sum score approach.

**Keywords:** dog; scale effects; environmental sensitivity.

**INTRODUCTION**

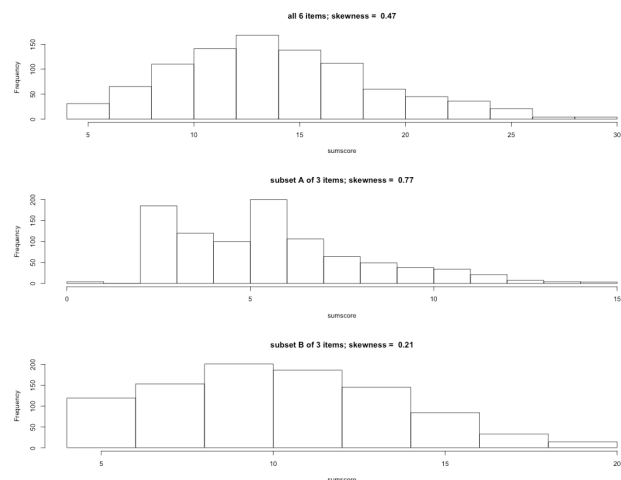
Non-additive effects like for instance genotype-environment interactions and genetically structured heterogeneity of residual variance are notoriously dependent on scale: many statistical non-additive phenomena disappear after a careful choice of transformation of the phenotypic values. Particularly for behavioural measures, scale is a delicate matter. Many behavioural tests change over time, for example by shortening the test and/or replacing items with new items. Such changes in the set of items in a test that assesses behaviour have a direct impact on the distribution of the sum score of a test (i.e., taking the sum of all items, the phenotypic score). Take for example a set of 6 items from the Dog Mentality Assessment (DMA; Svartberg and Forkman 2002) assessing fearful responses. Taking data from Irish soft-coated wheaten terriers, Figure 1 displays histograms of the summed score on all six items (top) and two different subsets of the same items. What items are included in a test directly affects the skewness of the distribution.

Especially if in a breeding setting different types of tests (say, an old version and a new one, or a long version and a short version) are used to measure the same personality trait, different subsets of the data can yield different conclusions regarding non-additive effects. This can be overcome if the distribution of test scores for the phenotype could be made independent of the particular items in the test.

Van den Berg et al. (2010) showed how the application of an Item Response Theory (IRT) model can incorporate the effects of items directly, thereby correcting inference for the specific characteristics of the items. With the scale being identified by the IRT measurement model, inference about genetic etiology can be applied to a scale that is always the same, regardless of how many and which

items are in a behavioural test. Schwabe and Van den Berg (in press) showed for human twin data that spurious results due to skewed sum score distribution can be prevented by incorporating IRT measurement models.

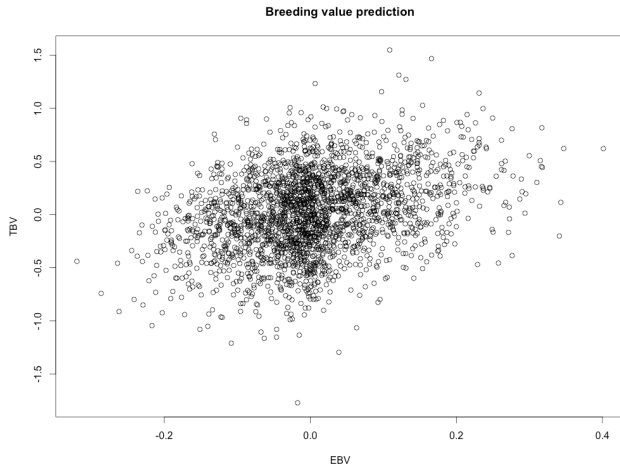
Here we present a Bayesian method for complex animal pedigrees assessing heterogeneity in residual variance that is linked to genotype and that incorporates an IRT measurement model for linking observed categorical data to an unobserved, latent phenotype. We apply the method to fearful behaviour in dogs.



**Figure 1.** Different shapes of the sum score distribution, depending on what items are in the test: all items (top; skewness=0.47), a subset A (skewness=0.77) or a subset B (skewness=0.21).

**MATERIALS AND METHODS**

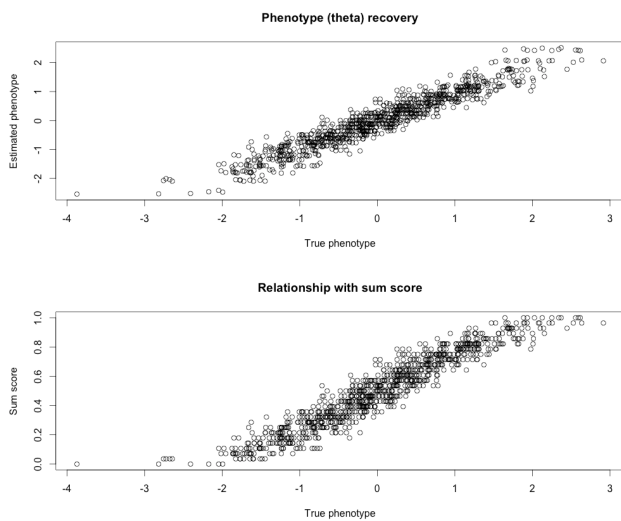
**Statistical analysis.** The dichotomous or dichotomized data point  $Y_{ij}$  for individual  $i$  and item  $j$  was modeled using a normal ogive item response theory model. Following Albert (1992), a data augmentation step was used by modeling an animal- and item-specific liability  $L$  that is a function of an unobserved phenotype  $\theta_i$  and an item characteristic  $b_j$ ,  $L_{ij} \sim N(\theta_i - \beta_j, 1)$ , and defining  $Y_{ij}=1$  if  $L_{ij} > 0$  and  $Y_{ij}=0$  otherwise. For latent phenotype  $\theta_i$  the usual univariate animal model was applied with  $\theta_i = X_i\beta + Z_i\mathbf{u} + e_i$ , with  $\mathbf{u} \sim N(0, A \cdot \text{var}_A)$  but  $\text{Var}(e_i) = \exp(\gamma_0 + \gamma_1 \cdot u_i)$ . We used relatively flat priors, for  $\beta$ ,  $\mathbf{b}$ ,  $\text{var}_A$  and  $\gamma$ . By determining the full conditional distributions, the model parameters can be sampled using a hybrid MCMC sampling scheme. Due to the application of a data augmentation step (Albert, 1992), the conditional distributions for  $\theta$  and  $\beta$  are of closed form and easily sampled from. The  $\gamma$  parameters can be sampled using a Langevin-Hastings step (Sorensen and Waagepetersen, 2003). For the Langevin-Hastings step, a scale factor of .005 was used.



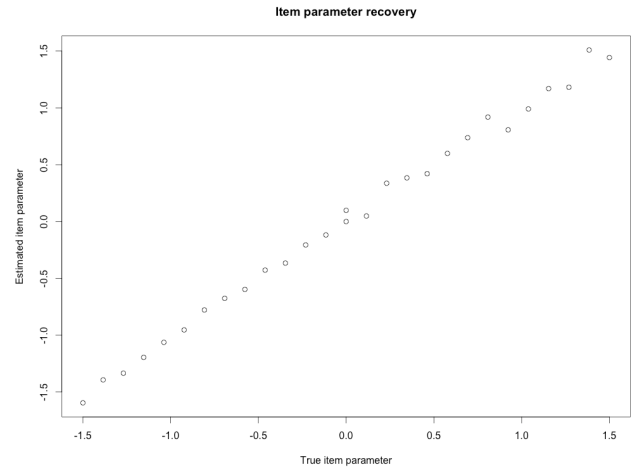
**Figure 2.** Recovery of simulated breeding values,  $h^2=15\%$ . EBV is posterior mean. Correlation is 0.34.

The vector of breeding values  $\mathbf{u}$  can be sampled with a Metropolis-Hastings step by having as proposal distribution the full conditional distribution of  $\mathbf{u}$  under the assumption that  $\gamma_1 = 0$ , where we follow García-Cortés and Sorensen (1996). The posterior analysis was based on 5000 iterations, with a burn-in of 1000 and a thinning factor of 2.

**Data.** Data came from the DMA. For 935 Irish soft-coated wheaten terriers there were complete behavioural data on a 6-item subscale for fearfulness (5-category items 17, 19, 20, 22, 23, and 24). The polytomous items were transformed into 24 dichotomous items using the procedure described by Van den Berg et al (2007). These dogs belonged to 514 full sib families and were progeny of 288 fathers and 377 mothers. Of these parents, 63 fathers and 95 mothers were themselves tested. For more than 55% of the tested dogs, at least one parent was tested. Untested relatives (1448 dogs) were included to facilitate estimation of genetic parameters. Maximum inbreeding coefficient was 0.25 (two dogs), with 0.0009 on average.



**Figure 3.** Recovery of simulated phenotypes (theta) and relationship with sum scores. Note the small variation in estimated phenotypic scores and sum scores for high and low true phenotypic scores.

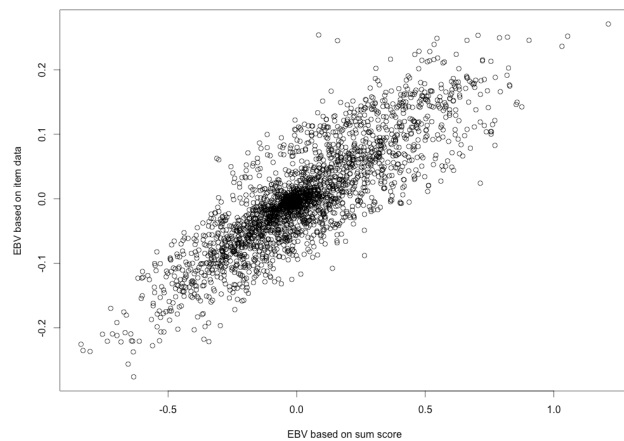


**Figure 4.** Recovery of simulated item parameters (beta).

**Simulation.** We also performed a simulation study in addition to the field data analysis. The pedigree and data structure were taken from Irish soft-coated wheaten terriers. Data were simulated with genetic variance 0.15 and environmental variance 0.85 with no heterogeneity. We used 28 dichotomous items, with beta parameters between -1.5 and 1.5.

## RESULTS AND DISCUSSION

**Simulation.** Figure 2 shows the recovery of simulated breeding values with similar data structure as in the dog sample, with 15% heritability, and no heterogeneity. Posterior mean for genetic variance was 0.16 (SD=0.05, for  $\gamma_0$  -0.04 (SD=0.08) and for  $\gamma_1$  -0.01 (SD=0.14). Figure 2 shows that extreme phenotypes are hard to estimate for any reasonable behavioural test (cf. Schwabe and Van den Berg, in press). Estimated phenotypes are known to be highly correlated with sum scores (Van den Berg et al 2007), so that with sum scores you also see that animals on the extremes of the true phenotypic scale tend to show little variation in sum scores. Ignoring this phenomenon can result in spurious



**Figure 5.** Estimated breeding values for fearfulness based on standardized sum score phenotypes (x-axis) and based on raw item data (y-axis),  $r = 0.85$ .

heterogeneity or masked true heterogeneity if sum scores are analysed. Here with the incorporation of the measurement model, the heterogeneity coefficient  $\gamma_1$  is close to its true value of zero. Figure 4 shows that item parameters  $\beta$  are recovered well.

**Fearfulness in dogs.** Table 1 shows the posterior means and standard deviations for the genetic variance and the parameters for the residual variance that is partly a function of genotypic value. The heterogeneity parameter  $\gamma_1$  is very different from 0 and negative, suggesting that at the latent phenotypic level sensitivity to environmental factors is lower for animals with high breeding values for fearfulness. In other words, dogs with a high genetic predisposition for fearfulness will tend to show fearful behaviour independent from the environment they are raised in and the treatment they have been exposed to. Conversely, sensitivity to environmental factors is higher for animals with low breeding values for fearfulness. Hence, dogs with a low genetic predisposition may anyway develop fearful behaviour if raised and kept under unfavourable conditions.

The results were compared with an analysis based on the sum of the 6 polytomous items (subsequently standardized to have unit variance). Genetic variance was estimated at 0.28 (SD=0.05),  $\gamma_0$  at -0.28 (SD=0.13) and  $\gamma_1$  at 0.28 (SD=0.11). Note the switching of the sign of parameter  $\gamma_1$ , which is now positive instead of negative. Figure 5 plots the EBVs for the two approaches.

**Table 1.** Posterior means (EAP) and standard deviations (SD) of model parameters for fearfulness in Irish soft-coated wheaten terriers.

Parameter	EAP	SD
Var_a	0.09	0.02
Gamma_0	1.24	0.09
Gamma_1	-0.41	0.16

## CONCLUSION

Estimation problems for high and low scoring animals on behavioural tests using sum scores can result in spurious heterogeneity of environmental variance or masking true heterogeneity at the latent level. Taking into account such measurement problems via an IRT model circumvents this problem and leads to unbiased estimates (cf. Schwabe and Van den Berg, in press). The method is particularly useful in breeding programs where different test versions are used and/or item data are missing at random. Here we showed that on the scale defined by a normal ogive IRT model, fearfulness in dogs is more sensitive to environmental factors for dogs with low breeding values for fearfulness. We also showed that analyzing sum scores can lead to the opposite conclusion (cf. Schwabe and Van den Berg, in press).

## LITERATURE CITED

- Albert, J. (1992). *J. Educ. Behav. Stat.* 17:251-269
- García-Cortés, L.A. and Sorensen, D. (1996). *Genet. Sel. Evol.* 28:121-126
- Schwabe, I. and Van den Berg, S.M. (in press). *Behav. Genet.*
- Sorensen, D. and Waagepetersen, R. (2003). *Genet. Res.* 82:207-222
- Svartberg, K. and Forkman, B. (2002). *Appl. Anim. Behav. Sc.* 79:133-155
- Van den Berg, S.M., Fikse, F., Arvelius, P. et al. (2010). *Proc 9th WCGALP*
- Van den Berg, Glas C.A.W. and Boomsma, D.I. (2007). *Behav. Genet.* 37:604-616: