AUTHOR	Barbara Class, Jon Brommer
TITLE	Contrasting multilevel relationships between behavior and body mass in blue tit nestlings
YEAR	2020
DOI	https://doi.org/10.1093/beheco/araa014
CITATION	Barbara Class, Jon Brommer, *Behavioral Ecology*, Volume 31, Issue 3, May/June 2020, Pages 702–710.

Title: Contrasting multi-level relationships between behavior and body mass in blue tit nestlings

Authors: Barbara Class[1], Jon Brommer[1]

Affiliations: Department of Biology, University of Turku, University Hill, 20014 Turku, Finland

Corresponding author: barbara.a.class@gmail.com

Funding

This work was supported by the Academy of Finland (grant number 289456 to J.E.B).

Acknowledgements

We thank landowners for giving us permission to work on their land and the people who helped us collecting data in the field. This work benefited from discussion with Kees Schreven. We also thank Alastair Wilson and three anonymous reviewers for providing thorough and constructive comments which helped us improve this manuscript.

Data accessibility: Analyses reported in this article can be reproduced using the data provided by Class and Brommer (2020).

Publication: Class B, Brommer JE. 2020. Contrasting multi-level relationships between behavior and body mass in blue tit nestlings. *Behavioral Ecology* 31, 702-710. DOI 10.1093/beheco/araa014

Abstract

Repeatable behaviors (i.e. animal personality) are pervasive in the animal kingdom and various mechanisms have been proposed to explain their existence. Genetic and non-genetic mechanisms, which can be equally important, predict correlations between behavior and body mass on different levels (e.g. genetic, environmental) of variation. We investigated multi-level relationships between body mass measured on weeks 1, 2, and 3 and three behavioral responses to handling, measured on week 3, and

which form a behavioral syndrome in wild blue tit nestlings. Using 7 years of data and quantitative genetic models, we find that all behaviors and body mass on week 3 are heritable (h^2= 0.18-0.23) and genetically correlated, whereas earlier body masses are not heritable. We also find evidence for environmental correlations between body masses and behaviors. Interestingly, these environmental correlations have different signs for early and late body masses. Altogether, these findings indicate genetic integration between body mass and behavior, and illustrate the impacts of early environmental factors and environmentally-mediated growth trajectory on behaviors expressed later in life. This study therefore suggest that the relationship between personality and body mass in developing individuals is due to various underlying mechanisms which can have opposing effects. Future research on the link between behavior and body mass would benefit from considering these multiple mechanisms simultaneously.

Keywords

Personality, behavioral syndrome, body mass, growth, heritability, genetic correlation, structural equation model, multivariate mixed model, *Cyanistes caeruleus*

Introduction

Animal personality is a pervasive phenomenon in nature, as evidenced by the great number of studies reporting repeatable (or heritable) behaviors and among-individual (or genetic) correlations between behaviors (i.e. behavioral syndromes) in wild animals. Animal personality research mainly aims at explaining among individual differences in behavior and a number of hypotheses involving genetic and/or phenotypic mechanisms have been proposed (Dingemanse and Réale 2013, Wolf et al. 2010). Both genetic and non-genetic mechanisms are likely to be equally important, as recently suggested by Dochtermann, Sih and Schwab (2015), who showed that additive genetic effects explain approximately half (52%) of among-individual variation in behavior, the remaining 48% being due to permanent environment effects. The latter include environmental factors that are maintained across measurements, of which an arguably major part are maternal and/or early environmental factors that have long-lasting effects on individuals' behavior through their impacts on its ontogeny (developmental plasticity, as

opposed to reversible plasticity). Roughly speaking, we hence can expect that genes that individuals inherit from their parents and the conditions they encounter early in life equally determine between-individual differences in behavior. Importantly, genes and environment can have opposing effects on behaviors, which is why genetic and environmental correlations between behaviors resulting from these effects do not always have similar signs (Dochtermann 2011).

Different hypotheses and theoretical models for the existence and maintenance of animal personality predict a close link between behavior and body mass. Both repeatable behaviors and body mass have well-documented associations with survival and reproductive success (but see Smith and Blumstein 2008, Moirón et al. 2020 for behaviors, and Wauters 1993, Profitt et al. 2008, Festa-Bianchet at al.1997, Linden et al. 1992, Monros et al. 2002 which show positive effects of juvenile body mass for survival in various bird and mammal species). Just as behavior, body mass is often heritable and strongly influenced by environmental conditions encountered by individuals. Therefore, body mass and behavior can be genetically and/or environmentally correlated.

Firstly, if genetically correlated, behavior and body mass can coevolve. For example, the pace-of life syndrome (POLS) hypothesis predicts genetic correlations between an array of behaviors (e.g. risk taking, parental care), physiological (e.g. metabolism, immune response) and life-history traits (e.g. growth rate, age at maturity), which can be generated by correlational selection or developmental constraints (Réale et al. 2010). Under this hypothesis, it may be more beneficial for individuals with a slow pace-of-life to have a slower growth rate, to be larger at sexual maturity (hence heavier) and to be less bold than individuals with a fast pace-of-life. On the other hand, body mass and behaviors can also be genetically correlated if underpinned by similar heritable mechanisms (e.g. metabolism, allometric constraints). For instance, positive genetic correlations between body mass and behaviors that enable food acquisition can be expected as heavier individuals have higher energy requirements (Mathot et al. 2018).

Secondly, phenotypic mechanisms can generate correlations between behavior and body mass. For example, consistent behavioral differences between individuals are predicted to arise and to be

correlated to body mass through the action of various positive or negative state-dependent feedback loops (Sih et al. 2015). However, such mechanisms remain empirically challenging to demonstrate and only apply to a certain type of study systems and behaviors (e.g. behaviors allowing food acquisition or predation avoidance). Behaviors which are not involved in such feedbacks may also be (directly or indirectly) related to body mass. For example, correlated plasticity to unmodelled environmental factors can affect both behavioral development/expression and mass, or can affect behavioral expression/development through effects on body mass, if behavior is mass-dependent. Reversible and developmental environmental plasticity can generate correlations between body mass and behavior in the short-term (behavior and body mass only correlate when measured at the same life stage) and in the long-term (behavior can be correlated to body mass at an earlier developmental stage which influences final mass), respectively.

Genetic and environmental relationships between body mass and behavior are therefore likely to arise as individuals develop. Importantly, body mass and growth are tightly related, but it remains unclear how they independently correlate with behavior. There is empirical support for correlations between growth rate and behavior (cf. meta-analysis in Royauté et al. 2018) and a few studies reported phenotypic associations between early-life body mass/birth weight and behaviors expressed later in life/personality in mammals (Rödel and von Holst 2009, Rödel and Meyer 2011, Hudson et al. 2011, Thomas et al. 2016), including humans (e.g. Hertz et al. 2013). Although a recent meta-analysis showed empirical evidence for among-individual correlation between body mass and behaviors (e.g. aggression or boldness, see Niemelä and Dingemanse 2018a), most reports of body mass-behavior association consist of phenotypic correlations (Niemelä and Dingemanse 2018a, b). At present there is hence a paucity of studies providing information on the role of environmental correlation (ecological mechanisms) or genetic correlations (evolutionary mechanisms) underlying phenotypic body mass-behavior correlations.

Because environmental and genetic relationships between body mass and behavior do not necessarily align, they need to be investigated using appropriate data and (co)variance partitioning approaches (Niemelä and Dingemanse 2018b). Blue tits (*Cyanistes caeruleus*) are an ideal study system in this

context as they often produce large families and their population characteristics facilitate the monitoring of a great number of them during the nesting period in natural conditions. Thanks to these features, and when data is available for a number of generations/breeding seasons, one can use the population's pedigree to partition the phenotypic covariance between behavior and body mass into different components (Kruuk 2004). Such components include covariance due to environmental conditions (e.g. territorial quality, clutch size) that differ between nests (common environment), parental effects, additive genetic effects, and differences among individuals that are not explained by the above-mentioned factors but can be caused, for example, by hatching asynchrony, genetic differences between siblings, or measurement error (residuals).

In the present paper, we investigate the multi-level relationship between early body masses measured 2, 9, and 16 days after hatching (D2, D9, D16) and three behavioral responses to handling measured on D16 in nestling blue tits. In this population, these behaviors are heritable and correlate on the additive genetic, common environment, and residual levels (Brommer and Kluen 2012) and all masses are expected to correlate positively with each other on all levels. Based on 7 years of phenotypic data collected in a pedigreed population and using a quantitative genetic model, we estimate the additive genetic, common environment and residual correlations between all six traits. For each estimated correlation matrix, we then compare 4 four structural equation models (SEM) which fit different hypotheses describing the potential relationships between body masses and a latent factor on which all behaviors load (Figure 1): i) under the "allometry hypothesis", all masses and behavior load on a single latent factor which is individuals' size; ii) under the "growth hypothesis", each mass is explained by earlier masses (growth sequence) and independently correlates with behavior; iii) under the "body mass" hypothesis, all masses are grouped under a single latent factor "body mass" which correlates with behavior; iv) under the "early vs. late mass hypothesis", early masses and late mass correlate independently with behavior while late mass is explained by early mass. Models were compared based on their AIC and uncertainty of the best model's estimates was obtained through parametric bootstrapping.

Material and methods

Data collection

Data used for these analyses was collected between 2012 and 2018 in a wild population of blue tits breeding in nest boxes in south-west Finland (Tammisaari, 60°01′N, 23°31′E). This population has been monitored yearly since 2003 during the breeding season (first broods from end of April to end of June), following a standard protocol for nest box-breeding passerines (Brommer and Kluen 2012). Nest boxes were visited weekly in May to assess laying dates, clutch sizes and estimate expected hatching dates. Nests from first broods were visited daily starting from their expected hatching date until at least one hatchling was observed (D0). Two days after the hatching day (D2), nestlings were weighed (using a scale with 0.1g precision) and their nails were clipped following unique combinations to allow their identification at later stages of development. Parents were caught and identified when nestlings were at least 5 days old. One week later (D9), nestlings were weighed and banded by putting a metal ring with a unique alphanumeric code on their left leg after their nail code was read. A few days before fledging (D16), nestlings were transferred all together in a large paper bag and various measurements of each nestling were taken following a fixed sequence (cf. Brommer and Kluen 2012). Firstly, each individual was held still on its back in the observer's palm. Stopwatch was started and the number of struggles during 10 seconds was counted. Docility was expressed as -1 time this number/second. Immediately after this 10 seconds assay, the time each bird took to take 30 breaths was measured twice using a stopwatch. Breath rate was calculated as 30 divided by the average of these two measures and expressed in number of breath/second. A higher breath rate reflects a higher stress response to handling (Carere et al. 2004). The bird's right tarsus and head-bill length were then measured using a digital sliding caliper (0.1mm accuracy) before measuring its wing and tail length using a ruler (1mm accuracy). During these morphometric measurements, the bird's aggressive behavior (struggling, flapping wings) was observed and a handling aggression score (1-5) was given to the bird. This score, which is 1 for a completely passive bird and 5 for a bird struggling continuously, reflects the time it takes for each bird to calm down during these measurements. Each nestling was then weighed using a Pesola spring balance (0.1g accuracy) and then placed in a second large paper bag where measured nestlings remained until the entire brood was processed and put back to its nest.

Pedigreed population

Phenotypic data was available for 5404 individuals, which were connected through a social pedigree based on social parenthood. The pruned pedigree, which retains only informative individuals holds record for 6205 individuals, 5464 maternities, 5107 paternities, 25107 full sibs, 43411 maternal sibs, 38543 paternal sibs, 18284 maternal half-sibs, 13416 paternal half-sibs, a mean family size of 10.8, a mean pairwise relatedness of 2.54e-3 and a maximum pedigree depth of 11. In this population, 11% to 22% of offspring produced annually were sired by extra-pair males (unpublished data). Based on simulations (Charmantier & Réale 2005), such level of error in paternity assignment is unlikely to cause substantial biases in quantitative genetic parameters when using the social pedigree.

Quantitative genetic analyses

Quantitative genetic analyses were carried out using animal models, which are mixed effects models that use the relatedness matrix derived from a population pedigree to estimate additive genetic (co) variance (Wilson et al. 2010). Univariate animal models assuming Gaussian distribution were run for each trait separately to estimate their variance components and their ratios to phenotypic variance. Then, a multivariate animal model was run for all six traits to estimate their correlation on various levels. In all models, brood identity, maternal identity, and additive genetic effects were fitted as random effects to estimate (co)variance due to common environment, maternal, and additive genetic effects while fixed effects included time of measurement in minutes and year as continuous and categorical covariates, respectively. For behavioral responses, fixed effects also included observer identity and handling order (continuous). In univariate models, box was fitted as an additional random effect to account for consistent differences between territories. Animal models were solved using Restricted Maximum Likelihood (REML), and implemented in ASReml-R version 3 (Butler et al. 2009; VSN International, Hemel Hempstead, U.K.). Statistical significance of fixed and random effects was tested using conditional Wald F tests and likelihood ratio tests (LRT) with one degree of freedom, respectively. Heritability (h^2) of each trait was calculated as the ratio V_A /V_P where V_P, the phenotypic variance, is defined as the sum of the REML estimates of additive genetic effects, maternal, common environment

effects and residuals (V_A , V_{PE}, and V_R respectively) and is conditional on the fixed-effect structure of the model. Correlations between pairs of traits on each level were calculated based on the corresponding covariance matrix estimated by the multivariate animal model. In this multivariate model, each response was corrected by the same fixed effects as in its corresponding univariate model. Random effects box and mother identities were not fitted in this model due to not being estimable for the former, and due to model convergence issues for the latter. Three 4-trait animal models including maternal effects were however fitted to verify that the relationships between each separate mass and behavior on other levels were consistent with the relationships found using the 6 trait animal model excluding maternal effects. Standard errors (SE) of variance ratios and correlations were approximated using the delta method (Fischer et al. 2004). Coefficients of variation (CV=sd*100/mean) were calculated for the different variance components in. All statistical analyses were performed in R (R development core team 2019). Residuals of all animal models were approximately normally distributed (Shapiro-Wilk test values >0.92, Figure S1).

Structural equation models

SEMs were used to investigate, on each level, different hypotheses for the relationships between behavior and body masses at different ages (Figure 1). SEMs have been previously used in behavioral studies to explore the structure of behavioral syndromes using predicted individual values derived from mixed models (Dochtermann & Jenkins 2007, Dingemanse et al. 2010). Here, each covariance matrix estimated by the multivariate model was converted into a correlation matrix, which was used as input data (cf. Class, Kluen and Brommer 2019, Moirón et al. 2019). In all SEMs variance of latent factors was fixed to 1. Because a correlation matrix was used as input data, the residual variance of each indicator (the variance unexplained by the latent factor) was fixed to 1 minus its squared factor loading. Each SEM was fitted in R using the package "lavaan" (Rosseel 2012). Sample size in these models was nominally set at 642 for the common environment level (number of broods) and 5404 (number of individuals) for the residual level and the all SEMs were compared using AIC. The sample size in a

SEM will not affect the inferred loadings or correlations between latent variables but can impact their uncertainty and the model AIC. We verified that the model rankings were similar if sample size was assumed to be lower.

Parametric bootstrap simulations were conducted to estimate median and 95% confidence intervals (CI) the model's loadings and assess model selection uncertainty. Because our findings indicated that the genetic covariance matrix was much reduced (see results), we focused on the common-environment and residuals covariances. Multivariate data for the 6 traits was simulated 1000 times using the inferred common environment and residual covariance matrices to generate a simulated dataset of the same dimension as the observations. Each simulated data was analyzed using a multivariate mixed model. At each iteration, and on each level, SEMs were run based on the estimated correlation matrices and ranked by AIC. Model selection uncertainty was assessed by calculating bootstrap selection rates (Lubke et al. 2017), which indicate whether model ranking is consistent to sampling variability. The selection rates have no a priori cut-off value for "significance", but instead provide an indication of model selection uncertainty (Lubke et al. 2017). For example, a selection rate of a SEM of 50% would indicate it is the top model in half the simulations, and 100% would suggest consistent support for this one SEM hypothesis over the others in all simulations. R code for performing SEMs and simulations are provided in Text S1.

Results

Univariate animal models

All traits have significant common environment variance and show no between-box variance (Table 1). We find significant additive genetic variance for behaviors, which all have a moderate heritability (0.16—0.22). In contrast, D9 and D16 masses have low heritabilities (0.01 and 0.12) and D2 mass heritability is not estimable. The proportion of phenotypic variance explained by common environment effects is moderate to low for behavioral traits (0.08—0.22) but higher for masses (0.29—0.43) and increasing between D9 and D16. Maternal variance overall represents a low proportion of phenotypic variance for behaviors (0.01 to 0.05) and masses (0.12) and only significantly differs from zero in D9

mass (and marginally for D16 mass). In contrast, residual variance accounts for more than half of the phenotypic variance in all behaviors (0.54—0.74). Coefficients of variation of the different variance components in masses show an increase in CV_A (0 to 3.35), and a decrease in CV_{CE} (15.09 to 6.44), CV_M (4.97 to 3.16) and CV_R (20.08 to 5.84), between D2 and D16 (Table 1). This phenomenon is likely due to compensatory growth (Figure S2) and the disappearance of the smallest broods and nestlings during the week separating two measurements (26% of the chicks and 20% of the broods measured on D2 die before D16; Table 1).

In all models, we find significant between-year differences (Tables S1-S6). Time of the day has a positive effect on masses on D2 and D9 but does not affect mass and behaviors measured on D16. We also find that observers differ in how they measure all behaviors and that handling order increases handling aggression and decreases breath rate but does not affect docility. Estimates of fixed and random effects as well as their test statistics and standard errors for each trait are reported in Tables S1-S6.

Multivariate animal models

On all levels, point estimates of correlations between handling aggression and docility and breath rate are negative (-0.26— -0.21 for HA-BR, -0.71— -0.30 for HA-docility) while estimates of the correlations between docility and breath rate are positive (0.10—0.45). In addition, estimates of correlations between masses are almost all positive on all levels (Table 2). The only exception is the genetic correlation between D2 and D9 masses which has a negative point estimate but also high uncertainty due to both masses having very low estimated V_A (Table S7). The high correlation estimates between subsequent mass measurements (0.54—0.79) indicate that rank orders of masses in nestlings and broods are mostly maintained from one week to the next.

Signs of the correlations between behaviors and D2 and D9 masses are consistent across all levels, although genetic correlations should be interpreted with caution given that both masses have a low (but estimable) V_A; On all levels, D2 and D9 masses appear to correlate positively with handling aggression (0.19—0.36) and negatively (-0.50— -0.08) with breath rate and docility. In contrast, the estimated

correlations between D16 mass and behaviors are inconsistent across levels; on the additive genetic level, D16 mass correlates positively with handling aggression (0.61) and negatively with breath rate and docility (-0.46,-0.37), whereas on the common environment and residual levels, D16 mass correlates negatively with handling aggression and docility (-0.04— -0.18) and positively with breath rate (0.05—0.25).

We tested the statistical significance of the correlations between D16 mass and behaviors on each level and whether these correlations differ across levels. To do so, we performed pairwise comparisons (LRT tests) between an unconstrained 4 trait animal model and similar models in which these three correlations were fixed to zero on one level or constrained to have similar values across two levels. These tests confirmed that genetic and common environment correlations between D16 mass and behaviors differ from zero (χ^2=9.21, df=3, p=0.03 and χ^2=26.57, df=3, p<0.001, respectively) and from each other (χ^2= 13.63 df=3, p=0.003), whereas residual correlations do not differ from zero (χ^2= 3.01, df=3, p=0.39), differ from common environment correlations (χ^2= 17.37 df=3, p=0.006) but not from genetic correlations(χ^2= 7.13, df=3, p=0.07). As a result of these inconsistent correlation signs across levels, phenotypic correlations between D16 mass and behaviors are low; point estimates are positive for handling aggression and breath rate (0.06, 0.03) and negative (-0.12) for docility (Table 2). Variances and covariances and their standard errors estimated by the multivariate animal model are reported in Table S7. Four-trait animal models fitting each separate mass and the 3 behaviors and including maternal effects resulted in genetic, common environment, and residual correlation matrices which were qualitatively similar to those obtained in the 6-trait model excluding maternal effects (Tables S8-S10).

Structural equation models

Because D2 and D9 masses have no or very low V_A (based on univariate animal models), interpretation of their estimated genetic correlations with each other and other traits are hampered. For this reason, the additive genetic correlation matrix was not considered further. We hence focused on common environment and residual correlation matrices for the SEM analysis. On both levels, the "early vs. late

mass" model (Fig. 1) is the best supported model (Tables 3 and 4). When setting a lower sample size for both levels, (e.g. arbitrarily set at 100 for the common environment level, 1000 for the residual level), rankings remained the same for the first and last-ranked models on the common environment level and for all models on the residual level (Table S11 and S12). Bootstrapped selection rates (for 1000 iterations) on the common environment level were 87% for the "early vs. late mass" model, 11% for the "growth" model and 2 % for the "body size" model, while the "early vs. late mass" model had 100% support on the residual level. Hence, the "early vs. late mass" was clearly the best-supported model on both levels. Bootstrapped estimates from this model indicate that early masses increase handling aggression but decrease docility and breath rate, while final mass has opposite effects on these behaviors (Figure 2).

Discussion

Using covariance partitioning and structural equation modeling, we investigated genetic and environmental relationships between body masses and a behavioral syndrome in developing blue tits. We show that that long-term (early-life) and short-term environmental factors have opposing effects on the relationship between body mass and behavior. In addition, we find that that the correlations between fledging mass and behavior have opposing signs on environmental and genetic levels. As a result, the phenotypic relationship observed between fledging mass and behavior is weak.

Firstly, we show that nestlings and broods that are heavier in their first two weeks express higher handling aggression and lower breath rate and docility in their third week. Because early body mass is mainly caused by environmental variation, with additive genetic effect essentially absent, our results suggest that environmental factors that nestlings experience during the first days after hatching have a lasting impact on their behavior at fledging. A similar long-term impact of early environment and mass on the ontogeny of behavior was shown in studies on laboratory rats, mice, European rabbits, and domestic cats (Rödel and von Holst 2009, Rödel and Meyer 2011, Hudson et al. 2011). Impacts of early nutrition on behavioral development have received more attention in the recent years and early-life food deprivation has been shown, mainly from lab experiments, to have long-lasting effects on adult behavior

in various vertebrate and invertebrate species (reviewed in Langenhof and Komdeur 2018). In zebra finches, female raised on low quality food are faster explorers (Krause et al. 2009) than females raised on high quality diet and males raised under low micronutrient diet are less bold than individuals raised under high micronutrient diets (Noguera et al. 2015). In great tits, food rationing in early life was shown to increase aggression in a line selected for high exploration and aggression (Carere et al. 2005). In wild great tits, nestlings provisioned with lower amounts of caterpillars by their parents were shown to exhibit a stronger stress response to handling and to be faster explorers later on (van Oers et al. 2015).

Secondly, we find that the environmental relationship between body mass and behaviors is positive until D9 and negative on D16, despite the overall maintenance of mass rankings. In this population, D9 corresponds to a critical developmental stage after which growth rates increase for lighter individuals catching up growth, or decrease for heavier individuals reaching their final mass. Such pattern is typical in species showing asymptotic growth. Although the mechanism involved remains unclear, this change of sign in the correlation between behavior and early vs. late body masses indicates that environmentally-mediated growth trajectories play a role in birds' behavioral development. There is some empirical evidence that growth trajectories (mediated by early nutrition) impact behavioral development. In particular, compensatory growth, despite immediate benefits, can incur costs that are paid later in life (Metcalfe and Monaghan 2001). For instance, compensatory growth has been shown to impair cognitive abilities (Fisher et al. 2006) and affect exploratory behavior (Krause and Naguib 2011) in zebra finches. The present study is one of the few studies documenting long-term impacts of early environment and growth trajectories on animal behavior in wild.

Thirdly, we find evidence for additive genetic covariance between D16 mass and a previously documented behavioral syndrome. In other words, genes increasing mass before fledging are associated with genes increasing handling aggression, and decreasing breath rate and docility in nestlings when they are handled. Because individuals express these behaviors in a stressful situation, this behavioral syndrome likely reflects how individuals cope with stress (i.e. coping styles, Koolhaas et al. 1999). We can thus expect higher levels of expression of this behavioral syndrome (i.e. higher handling aggression, lower breath rate and docility) to be associated with a proactive coping style and a fast pace-of-life. Our

findings that higher levels of expression of this behavioral syndrome are genetically associated with higher fledging masses hence seem to contrast with the POLS prediction that smaller individuals show more proactive behaviors and are less stressed. Nevertheless, this study does not allow ruling out the possibility of coevolution between these behaviors and body mass, as a result of correlational selection. Interestingly, and although the allometry hypothesis is not supported, the negative relationship that we find between body mass and breath rate aligns with the well-documented allometric relationship between body mass and respiratory variables across species (e.g. Frappell et al. 2001).

Regardless of its mechanistic cause (i.e. pleiotropy or linkage disequilibrium), the genetic association between behaviors and D16 body mass indicates that these traits will respond non-independently to selection. Although it remains unclear whether and how these behavioral responses to handling influence post-fledging survival, mass at fledging is known to be a strong predictor of survival in small passerines (e.g., Perrins, 1965, Tinbergen and Boerlijst 1990, Linden et al.1992, Monrós et al. 2002, Radersma et al. 2015). Therefore selection for a higher mass should result in selection for a higher expression of this behavioral syndrome. Importantly, the behavior of an individual at fledging is genetically correlated to its behavior as a reproductive adult (Class and Brommer 2015), which was shown to influence adult survival and reproductive success (Class et al. 2014). As a result, some or several of the behaviors as well as the genetically correlated mass measured at fledging are likely under selection. Quantification of multivariate selection in combination with the here documented additive genetic covariances for behaviors and mass at fledging is required to investigate whether maintenance of variation in behaviors is facilitated when taking a multivariate perspective on the evolutionary dynamics of this behavioral syndrome or not.

Theoretical models proposed adaptive integration of state (e.g. body mass, size) and behavior as a possible mechanism explaining the existence of animal personality (e.g. Wolf et al. 2007) but still lack empirical support. Using a combination of multivariate mixed models and a SEM approach similar to what was done in this study, Moirón et al. (2019) recently found evidence for phenotypic integration between risk-taking behaviors and body size in adult great tits. The present study provides further evidence for such integration on the genetic level. In addition, our partitioning of phenotypic variance

into additive genetic and common environment effects reveals complex relationships between body mass and behavior, which change as individuals grow and do not align across levels. Importantly, despite their strong genetic correlation, body mass and behavior are only weakly correlated on the phenotypic level, which cautions against taking the phenotypic gambit (Cheverud 1988, Dochtermann 2011). Recent meta-analysis (Niemelä and Dingemanse 2018) revealed that among-individual correlations between intrinsic state variables (e.g. body size and mass) and behaviors (e.g. aggression or boldness) are weakly positive and that individual variation in states only explains about 5% of phenotypic variation in these behaviors. In the light of our findings, it is possible that high genetic correlations exist between states and behaviors but are being masked by other sources of (co)variation. Because quantitative genetic models allow partitioning such (co) variation further, they constitute valuable tools to study the coevolution between state and behavior and hence to understand better the existence of animal personality.

This study provided insights into the relationship between handling behaviors and body mass in developing wild birds. Our findings demonstrated the importance of early environmental factors and environmentally-mediated growth trajectory for behavioral development and provided evidence for a genetic association between this behavioral syndrome and fledging mass. We also show non-aligning genetic and environmental correlations between fledging body mass and behavior which result in a weak phenotypic correlation. Altogether, these results exemplify the complexity of the relationship between behavior and body mass, shaped by genetic, long-term, and short-term environmental factors. Such complex and interesting relationship would have been missed, had we ignored earlier masses and not partitioned phenotypic (co)variances. Research on the link between behavior and body mass under natural environmental conditions can advance our understanding of animal personalities and would benefit from considering multiple underlying mechanisms simultaneously.

References

Brommer, JE and Kluen, E. 2012. Exploring the genetics of nestling personality traits in a wild passerine bird: testing the phenotypic gambit. *Ecol Evol.* **2**:3032–3044.

Butler, DG, Cullis, BR, Gilmour, AR, Gogel, BJ. 2009. ASReml-R reference manual.

Carere, C and van Oers, K. 2004. Shy and bold great tits (Parus major): body temperature and breath rate in response to handling stress. *Physiol. Behav.* **82**: 905–912.

Charmantier, A and Réale D. 2005. How do misassigned paternities affect the estimation of heritability in the wild ? *Mol Ecol.* **14**:2839–2850.

Cheverud, JM. 1988. A Comparison of Genetic and Phenotypic Correlations. *Evolution.* **42**: 958–968.

Class, B, Kluen, E, Brommer, JE. 2014. Evolutionary quantitative genetics of behavioral responses to handling in a wild passerine. *Ecol. Evol.* **4**: 427–440.

Class, B and Brommer, JE. 2015. A strong genetic correlation underlying a behavioural syndrome disappears during development because of genotype-age interactions. *Proc. R. Soc. B Biol. Sci.* **282**: 20142777.

Class, B, Kluen, E, Brommer, JE. 2019. Tail colour signals performance in blue tit nestlings. *J. Evol. Biol.* **32**: 913–920.

Class, B and Brommer, JE. 2020. Data from: Contrasting multi-level relationships between behavior and body mass in blue tit nestlings. Behavioral Ecology. https://doi.org/10.5061/dryad.9ghx3ffdj

Dingemanse NJ, Dochtermann N, Wright J. 2010. A method for exploring the structure of behavioural syndromes to allow formal comparison within and between data sets. *Anim Behav.* **79**:439–450.

Dingemanse, NJ and Réale, D. 2013. What is the evidence that natural selection maintains variation in animal personalities? In: *Animal personalities : behavior, physiology, and evolution* (C. Carere and D. Maestripieri, eds), pp. 201–220. The University of Chicago Press, Chicago.

Dochtermann NA, Jenkins SH. 2007. Behavioural syndromes in Merriam's kangaroo rats (Dipodomys merriami): A test of competing hypotheses. *Proc R Soc B Biol Sci.* **274**:2343–2349.

Dochtermann NA. 2011. Testing Cheverud's conjecture for behavioral correlations and behavioral syndromes. *Evolution*. **65**:1814–1820.

Dochtermann, NA, Schwab, T and Sih, A. 2015. The contribution of additive genetic variation to personality variation: heritability of personality. *Proc. R. Soc. B Biol. Sci.* **282**: 20142201.

Festa-Bianchet, M, Jorgenson, JT, Bérubé, CH, Portier, C and Wishart, WD. 1997. Body mass and survival of bighorn sheep. *Can. J. Zool.* **75**: 1372–1379.

Fischer, TM, Gilmour, AR, Van Der Werf, JH. 2004. Computing approximate standard errors for genetic parameters derived from random regression models fitted by average information REML. *Genet Sel Evol.* **36**:363–369.

Fisher, MO, Nager, RG and Monaghan, P. 2006. Compensatory growth impairs adult cognitive performance. *PLoS Biol.* **4**: 1462–1466.

Frappell, PB, Hinds, DS, Boggs, DF. 2001. Scaling of respiratory variables and the breathing pattern in birds: an allometric and phylogenetic approach. *Physiological and Biochemical Zoology*, **74**, 75-89.

Hertz CL, Mathiasen R, Hansen BM, Mortensen EL, Greisen G. 2013. Personality in Adults Who Were Born Very Preterm. *PLoS One.* **8**: e66881.

Hudson, R, Bautista, A, Reyes-Meza, V, Montor, JM and Rödel, HG. 2011. The effect of siblings on early development: A potential contributor to personality differences in mammals. *Dev. Psychobiol.* **53**: 564–574.

Koolhaas JM, Korte SM, De Boer SF, Van Der Vegt BJ, Van Reenen CG, Hopster H, De Jong IC, Ruis MAW, Blokhuis HJ. 1999. Coping styles in animals: current status in behavior and stress-physiology. *Neurosci Biobehav Rev.* **23**:925–935.

Krause, ET and Naguib, M. 2011. Compensatory growth affects exploratory behaviour in zebra finches, Taeniopygia guttata. *Anim. Behav.* **81**: 1295–1300.

Kruuk, LEB. 2004. Estimating genetic parameters in natural populations using the "animal model." *Philos. Trans. R. Soc. B Biol. Sci.* **359**: 873–890.

Langenhof, MR, and Komdeur, J. 2018. Why and how the early-life environment affects development of coping behaviours. *Behav. Ecol. Sociobiol.* **72**: 34. Behavioral Ecology and Sociobiology.

Linden, M, Gustafsson, L and Part, T. 1992. Selection of fledging mass in the collared flycatcher and the great tit. *Ecology* **73**: 336–343.

Lubke GH, Campbell I, McArtor D, Miller P, Luningham J, van den Berg SM. 2017. Assessing Model Selection Uncertainty Using a Bootstrap Approach: An update. *Struct Equ Model.* **24**:230–245.

Luttbeg, B and Sih, A. 2010. Risk, resources and state-dependent adaptive behavioural syndromes. *Philos. Trans. R. Soc. B* **365**: 3977–3990.

Mathot, K. J., N. J. Dingemanse, and S. Nakagawa. 2019. The covariance between metabolic rate and behaviour varies across behaviours and thermal types: meta-analytic insights. Biol. Rev. 94:1056–1074.

Metcalfe, NB and Monaghan, P. 2001. Compensation for a bad start: Grow now, pay later? *Trends Ecol. Evol.* **16**: 254–260.

Moirón , M, Araya-Ajoy, YG, Mathot, KJ, Mouchet, A and Dingemanse, NJ. 2019. Functional relations between body mass and risk-taking behavior in wild great tits. *Behav. Ecol.* **30**: 617–623.

Moirón, M., K. L. Laskowski, and P. T. Niemelä. 2020. Individual differences in behaviour explain variation in survival: a meta-analysis. *Ecol. Lett.* **23**:399–408.

Monrós, JS, Belda, EJ and Barba, E. 2002. Post-fledging survival of individual great tits : the effect of hatching date and fledging mass. *Oikos* **99**: 481–488.

Niemelä, PT and Dingemanse, NJ. 2018. Meta-analysis reveals weak associations between intrinsic state and personality. *Proc. R. Soc. B Biol. Sci.* **285**.

Niemelä, PT and Dingemanse NJ. 2018. On the usage of single measurements in behavioural ecology research on individual differences. *Anim Behav.* **145**:99–105.

Noguera, JC, Metcalfe, NB, Surai, PF and Monaghan, P. 2015. Are you what you eat? Micronutritional deficiencies during development influence adult personality-related traits. *Anim. Behav.* **101**: 129–140.

Perrins, CM. 1965. Population Fluctuations and Clutch-Size in the Great Tit, Parus major L. *J. Anim. Ecol.* **34**: 601–647.

Proffitt, KM, Garrott, RA and Rotella, JJ. 2008. Long-term evaluation of body mass at weaning and postweaning survival rates of Weddell seals in Erebus Bay, Antarctica. *Mar. Mammal Sci.* **24**: 677–689.

R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Radersma, R, Komdeur, J and Tinbergen, JM. 2015. Early morning fledging improves recruitment in Great Tits Parus major. *Ibis* **157**: 351–355.

Réale, D, Garant, D, Humphries, MM, Bergeron, P, Careau, V and Montiglio, P. 2010. Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philos. Trans. R. Soc. B* **365**: 4051–4063.

Rödel, HG and Meyer, S. 2011. Early development influences ontogeny of personality types in young laboratory rats. *Dev. Psychobiol.* **53**: 601–613.

Rödel, HG and von Holst, D. 2009. Features of the early juvenile development predict competitive performance in male European rabbits. *Physiol. Behav.* **97**: 495–502.

Rosseel, Y. 2012. lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, **48**:1–36.

Royauté, R., M. A. Berdal, C. R. Garrison, and N. A. Dochtermann. 2018. Paceless life? A meta-analysis of the pace-of-life syndrome hypothesis. *Behav. Ecol. Sociobiol*. **72**:64.

Sih, A, Mathot, KJ, Moirón, M, Montiglio, P, Wolf, M and Dingemanse, NJ. 2015. Animal personality and state – behaviour feedbacks : a review and guide for empiricists. *Trends Ecol. Evol.* **30**: 50–60.

Smith, B. R., and D. T. Blumstein. 2008. Fitness consequences of personality: a meta-analysis. *Behav. Ecol.* **19**:448–455.

Thomas P, Herrel A, Hardy I, Aujard F, Pouydebat E. 2016. Exploration Behavior and Morphology are Correlated in Captive Gray Mouse Lemurs (Microcebus murinus). *Int J Primatol.* **37**:405–415.

Tinbergen, J and Boerlijst, M. 1990. Nestling Weight and Survival in Individual Great Tits (Parus major). *J. Anim. Ecol.* **59**: 1113–1127.

van Oers, K, Kohn, GM, Hinde, CA and Naguib, M. 2015. Parental food provisioning is related to nestling stress response in wild great tit nestlings: Implications for the development of personality. *Front. Zool.* **12**: S10.

Wauters, L, Bijnens, L and Dhondt, AA. 1993. Body Mass at Weaning and Juvenile Recruitment in the Red Squirrel. *J. Anim. Ecol.* **62**: 280–286.

Wilson, AJ, Reale, D, Clements, MN, Morrissey, MM, Postma, E, Walling, CA, *et al.* 2010. An ecologist's guide to the animal model. *J. Anim. Ecol.* **79**: 13–26.

Wolf, M and Weissing, FJ. 2010. An explanatory framework for adaptive personality differences. *Philos. Trans. R. Soc. B* **365**: 3959–3968.

Figure 1: Four alternative structural equation models, which can describe the relationship between behaviors and body mass. Continuous arrows represent the loadings of variables (rectangular shape) on latent factors (oval shape). In all models, handling aggression, breath rate, and docility (HA, BR, Doc respectively) load on a latent behavioral factor. Double-headed spotted arrows represent the correlations between behaviors and body masses. The one headed spotted arrow in "early vs. late mass model" represents the causal relationship between the latent factor "early mass" and D16 mass.

Figure 2: Median coefficients of the best supported model ("early vs. late mass") on the common environmental (A) and residual (B) levels and their 95% confidence intervals estimated by parametric bootstrapping (1100 iterations). Sample sizes for the common environment and residual levels were 1031 and 1051, respectively, after non-converging models were excluded.

Table1: Sample sizes, mean (and standard deviation), animal model variance estimates (V_A, V_{CE}, V_M, V_R, V_{box}), their ratios to VP (and their standard error) for all six traits (HA = handling aggression, BR= breath rate, Doc.= docility). Variance estimates for which LRT returned a p-value <0.05 (or <0.10) are printed in bold (and bold italics), respectively. Coefficients of variation are printed in italics below variance estimates. Variance components and ratios that cannot be estimated are noted n.e..

Trait	n nestlings	n broods	Mean	V_A	V_{CE}	V_M	V_R	V_{box}	h^2	V_{CE}/V_P	V_M/V_P	V_R/V_P
D2 mass	5268	627	1.86 (0.49)	n.e.	**0.079 (0.008)** *15.9*	0.009 (0.007) *4.97*	0.140 (0.003) *20.08*	n.e.	n.e.	0.35 (0.03)	0.03 (0.03)	0.62 (0.02)
D9 mass	4577	557	8.60 (1.33)	0.018 (0.109) *1.56*	**0.472 (0.066)** *7.99*	**0.194 (0.073)** *5.12*	0.934 (0.059) *11.24*	n.e.	0.01 (0.07)	0.29 (0.04)	0.12 (0.04)	0.58 (0.04)
D16 mass	4119	526	11.40 (1.16)	0.146 (0.116) *3.35*	**0.539 (0.070)** *6.44*	***0.130 (0.070)*** *3.16*	0.442 (0.060) *5.83*	n.e.	0.12 (0.09)	0.43 (0.05)	0.10 (0.05)	0.35 (0.05)
HA	4110	526	2.96 (1.27)	**0.295 (0.099)** *18.38*	**0.283 (0.048)** *17.99*	0.009 (0.047) *3.23*	0.783 (0.054) *29.91*	n.e.	0.22 (0.07)	0.21 (0.03)	0.01 (0.03)	0.57 (0.05)
BR	4114	526	2.02 (0.43)	**0.030 (0.012)** *8.59*	**0.036 (0.006)** *9.41*	0.009 (0.006) *4.58*	0.086 (0.007) *14.56*	n.e.	0.19 (0.08)	0.22 (0.04)	0.05 (0.04)	0.54 (0.05)
Doc.	4102	525	-0.21 (0.18)	**0.005 (0.002)** *32.67*	**0.002 (0.001)** *22.62*	0.001 (0.001) *12.81*	0.022 (0.001) *69.89*	0.001 (4.90E-4)	0.16 (0.06)	0.08 (0.03)	0.02 (0.03)	0.74 (0.04)

Table 2: Correlations (and standard errors) estimated by the multivariate animal model on the additive genetic, common environment, residual, and phenotypic levels. HA and BR denote handling aggression and breath rate. Correlations between masses and behaviors are highlighted. Statistical significance of correlations was not tested individually but estimates that are higher than two times their standard error are printed in bold.

Level		D2 mass	D9 mass	D16 mass	HA	BR
Additive genetic	D9 mass	-0.49 (1.55)				
	D16 mass	0.04 (0.72)	0.63 (0.32)			
	HA	0.26 (0.49)	0.36 (0.34)	**0.61 (0.27)**		
	BR	-0.08 (0.51)	-0.50 (0.38)	-0.46 (0.29)	-0.26 (0.17)	
	Docility	-0.20 (0.48)	-0.38 (0.33)	-0.37 (0.24)	**-0.71 (0.11)**	**0.45 (0.16)**
Common environment	D9 mass	**0.54 (0.05)**				
	D16 mass	0.10 (0.07)	**0.59 (0.05)**			
	HA	**0.29 (0.08)**	**0.29 (0.08)**	-0.11 (0.09)		
	BR	**-0.18 (0.08)**	-0.10 (0.09)	**0.25 (0.08)**	**-0.24 (0.09)**	
	Docility	**-0.25 (0.10)**	**-0.37 (0.10)**	-0.18 (0.10)	**-0.41 (0.10)**	0.10 (0.12)
Residual	D9 mass	**0.79 (0.02)**				
	D16 mass	**0.38 (0.05)**	**0.68 (0.04)**			
	HA	**0.21 (0.04)**	**0.19 (0.04)**	-0.02 (0.06)		
	BR	**-0.15 (0.04)**	**-0.12 (0.05)**	0.05 (0.07)	**-0.21 (0.05)**	
	Docility	**-0.14 (0.03)**	**-0.14 (0.04)**	-0.04 (0.05)	**-0.30 (0.03)**	**0.22 (0.04)**
Phenotypic	D9 mass	**0.64 (0.01)**				
	D16 mass	**0.21 (0.01)**	**0.61 (0.01)**			
	HA	**0.22 (0.02)**	**0.21 (0.02)**	**0.06 (0.02)**		

	BR	**-0.14 (0.02)**	**-0.13 (0.02)**	0.03 (0.02)	**-0.22 (0.01)**	
	Docility	**-0.16 (0.02)**	**-0.19 (0.02)**	**-0.12 (0.02)**	**-0.40 (0.01)**	**0.24 (0.01)**

Table 3: Model comparison on the common environment level. SEM models are drawn in Figure 1.

Model	ΔAIC	AIC	n.parameters
Early vs. late mass model	0	10082.44	16
Growth model	22.39	10104.82	15
Body mass model	24.56	10106.99	15
Allometry model	144.11	10226.55	13

Table 4: Model comparison on the residual level. SEM models are drawn in Figure 1.

Model	ΔAIC	AIC	n.parameters
Early vs. late mass model	0	81537.75	16
Body mass model	94.05	81631.79	15
Allometry model	461.40	81999.15	13
Growth model	880.63	82418.38	15

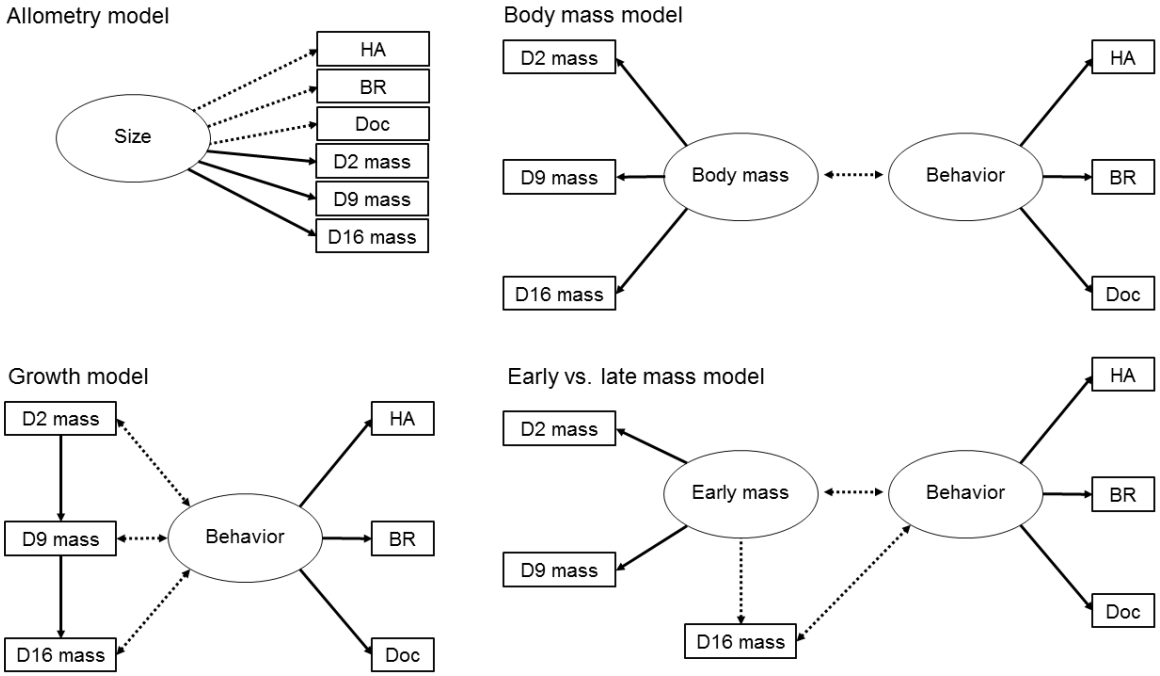
Allometry model
Size
HA
BR
Doc
D2 mass
D9 mass
D16 mass
Body mass model
D2 mass
D9 mass
D16 mass
Body mass
Behavior
HA
BR
Doc
Growth model
D2 mass
D9 mass
D16 mass
Behavior
HA
BR
Doc
Early vs. late mass model
D2 mass
D9 mass
D16 mass
Early mass
Behavior
HA
BR
Doc

Figure 1

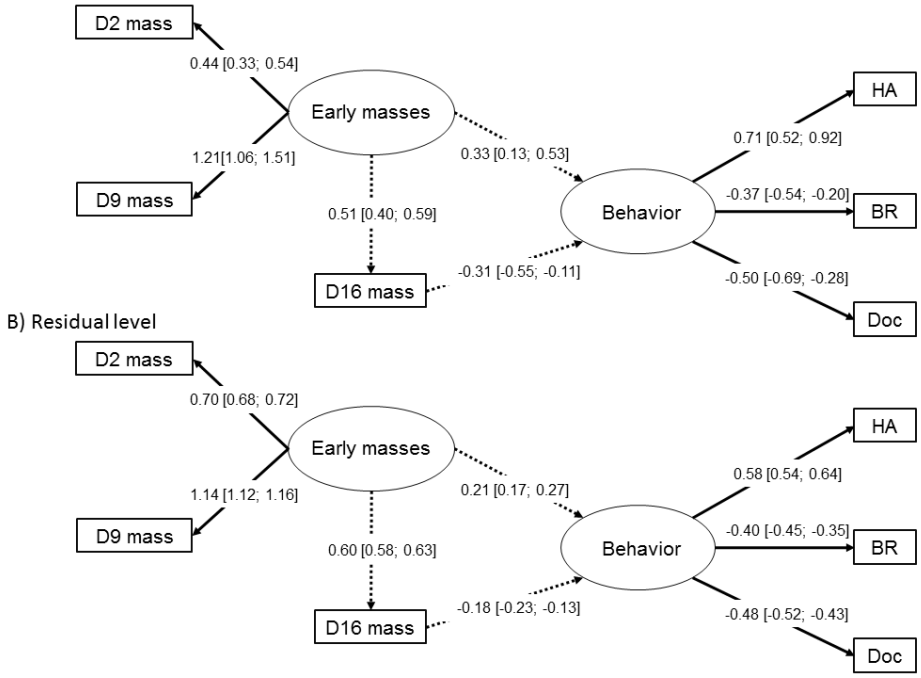
D2 mass
0.44 [0.33; 0.54]
Early masses
1.21[1.06; 1.51]
D9 mass
0.51 [0.40; 0.59]
D16 mass
0.33 [0.13; 0.53]
-0.31 [-0.55; -0.11]
Behavior
0.71 [0.52; 0.92]
-0.37 [-0.54; -0.20]
-0.50 [-0.69; -0.28]
HA
BR
Doc
B) Residual level
D2 mass
0.70 [0.68; 0.72]
Early masses
1.14 [1.12; 1.16]
D9 mass
0.60 [0.58; 0.63]
D16 mass
0.21 [0.17; 0.27]
-0.18 [-0.23; -0.13]
Behavior
0.58 [0.54; 0.64]
-0.40 [-0.45; -0.35]
-0.48 [-0.52; -0.43]
HA
BR
Doc

Figure 2