Genetics of social interactions in laying hens:

improving survival and productivity

Katrijn Peeters

Thesis committee

Promotor
Prof. Dr J.A.M. van Arendonk
Professor of Animal Breeding and Genetics
Wageningen University

Co-promotor
Dr P. Bijma
Assistant professor, Animal Breeding and Genomics Centre
Wageningen University

Other members
Prof. Dr M.C.M. de Jong, Wageningen University
Dr J.D. Hadfield, University of Edinburgh, UK
Prof. Dr M.J. Wade, Indiana University, USA
Dr B. Ask, Pig Research Centre, Denmark

This research was conducted under the auspices of the Graduate School of Wageningen Institute of Animal Sciences (WIAS)

Genetics of social interactions in laying hens:

improving survival and productivity

Katrijn Peeters

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 26 October 2015
at 4 p.m. in the Aula.

Katrijn Peeters
Genetics of social interactions in laying hens: improving survival and productivity,
168 pages.

PhD thesis, Wageningen University, Wageningen, NL (2015)
With references, with summary in English

ISBN 978-94-6257-547-9

Abstract

Through social interactions, individuals affect one another's phenotype. The heritable effect of an animal on its own phenotype is referred to as a direct genetic effect (DGE), while the heritable effect of an animal on the phenotype of a conspecific is referred to as an indirect genetic effect (IGE). Both DGEs and IGEs determine a population's potential to respond to selection, *i.e.* the total genetic variance. This thesis focusses on the genetic architecture of survival time in laying hens showing feather pecking and cannibalistic behaviour, a well-known social interaction trait. DGEs and IGEs for survival time were estimated in purebred and crossbred laying hens. Unrelated birds of the same line or cross were kept in groups of four. Beaks were kept intact. Results showed that IGEs contribute around half of the total genetic variance in purebreds and the majority of the total genetic variance in crossbreds (up to 87%). The direct-indirect genetic correlations were close to zero in purebreds and moderately to highly negative in crossbreds. Consequently, unlike purebreds, crossbreds would fail to respond positively to mass selection. To ensure positive response to selection, animals should be selected based on their total breeding value. Moreover, increased response to selection can be obtained when including genotypic information. With genomics, the accuracy of estimated breeding values increased with 20 up to 110%, showing the added value of genotypic information.

In addition, the genetic correlation between survival time (individual data) and early egg production (pooled data) was calculated. Results showed that, unlike for individual data, pooled data cannot be used to estimate DGEs and IGEs. However, pooled data can be used to estimate total genetic effects. The default bivariate model did not allow all non-genetic correlations between both traits to be fitted and, therefore, resulted in biased genetic parameter estimates. When this issue was resolved, the genetic correlation between survival time and early egg production was slightly negative (-0.09), but not significantly different from zero.

Finally, the interpretation of T^2 as a measure of inheritance for social interaction traits was discussed. T^2 expresses the total genetic variance relative to the phenotypic variance. Throughout this thesis it became clear that, for social interaction traits, the level of data collection (individual *vs* pooled data) and the within-group relatedness affects the phenotypic variance and, consequently, affects T^2. Therefore, T^2 can differ between experimental set-ups, even though the underlying genetic parameters are the same. This is undesirable for the comparison of studies. For survival time, a 30 up to 40% decrease in T^2 was observed when using pooled data instead of individual data. This illustrates that T^2, as a measure of inheritance, should be used with care.

Contents

1

General introduction

1.1 Social interactions

Social interactions among individuals are widespread, both in wild and domestic populations (Frank, 2007). By interacting, individuals might positively or negatively affect one another's phenotype. The effect of an animal on its own phenotype is referred to as a direct effect, while the effect of an animal on the phenotype of a conspecific is referred to as an indirect effect (Moore *et al.*, 1997). For example, for survival in group-housed laying hens, the direct effect is the effect of an animal on its own survival, while the indirect effect is the effect of an animal on the survival of its group mates. Both the direct and indirect effect can have a heritable component. Indirect genetic effects are also referred to as associative or social genetic effects (Dickerson, 1947; Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007). Traditionally, geneticists only modelled direct genetic effects, while the indirect genetic effects remained ignored. The only exception is maternal genetic effects, which have been studied extensively (Dickerson, 1947; Willham, 1963; Cheverud, 1984; Kirkpatrick and Lande 1989; Koerhuis and Thompson 1997; Mousseau and Fox 1998; Eaglen and Bijma 2009; Bouwman *et al.*, 2010). Maternal genetic effects are a specific kind of indirect genetic effects, where the indirect effect of a mother on the phenotype of her offspring has a genetic component (Dickerson, 1947; Willham, 1963).

In livestock, indirect genetic effects have been estimated for a variety of traits and species, *e.g.* survival in laying hens, body weight in quail, growth in pigs, fin erosion in fish, *etc.* (Ellen *et al.*, 2008; Muir *et al.*, 2013; Chen *et al.*, 2008; Canario *et al.*, 2010; Bouwman *et al.*, 2010; Nielsen *et al.*, 2014). Together with theoretical work (Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007), these studies show that a large part of the heritable variation can remain undetected if indirect genetic effects are not modelled. One of the most extreme examples currently reported is body weight in quail, where 88% of the heritable variation remained undetected with a classic direct animal model (Muir *et al.*, 2013). In addition, if direct and indirect genetic effects are negatively correlated, selection based on classic direct breeding values can result in negative response to selection (Griffing, 1967; Muir *et al.*, 2013).

In conclusion, to maximize response to selection for social interaction traits, indirect genetic effects should be accounted for.

1.2 Direct-indirect genetic models

This section briefly summarizes the quantitative genetic principles of social interaction traits. Table 1.1 gives an overview of the symbols that will be used frequently throughout this thesis.

Table 1.1 Notation key

Symbol	Meaning
i	Focal individual
j	Group mate
DGE; A_D	Direct genetic effect; direct breeding value
IGE; A_I	Indirect genetic effect; indirect breeding value
TGE; A_T	Total genetic effect; total breeding value
E_D	Direct environmental effect
E_I	Indirect environmental effect
$\sigma^2_{A_D}$	Direct genetic variance
$\sigma_{A_{DI}}$	Direct-indirect genetic covariance
$\sigma^2_{A_I}$	Indirect genetic variance
$\sigma^2_{A_T}$	Total genetic variance
r_{DI}	Direct-indirect genetic correlation
h^2	Direct heritable variance relative to phenotypic variance
T^2	Total heritable variance relative to phenotypic variance

For non-social interaction traits, an animal's phenotype (P) depends on its direct genetic (A_D) and environmental (E_D) effect (Fisher, 1918; Lynch and Walsh, 1998):

$$P_i = A_{D_i} + E_{D_i}. \quad [1.1]$$

For social interaction traits, an animal's phenotype depends on the direct genetic (A_D) and environmental (E_D) effect of the animal itself (i), and the indirect genetic (A_I) and environmental (E_I) effect of each of its group mates (j) (Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007):

$$P_i = A_{D_i} + \sum_{i \neq j}^{n-1} A_{I_j} + E_{D_i} + \sum_{i \neq j}^{n-1} E_{I_j}, \quad [1.2]$$

where n is the total number of group members (Figure 1.1). Consequently, each individual has $n - 1$ group mates. Note that an animal's A_D is expressed in the phenotype of the animal itself, while an animal's A_I is expressed in the phenotype of each of its group mates (Equation 1.2; Figure 1.2).

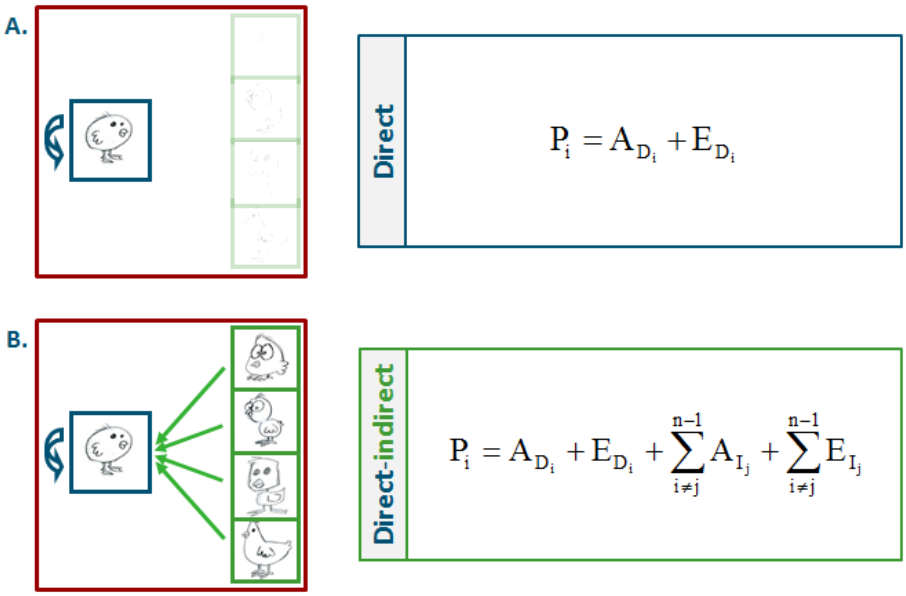

Figure 1.1 Genetic and environmental components determining the phenotype of non-social (A) and social (B) interaction traits

To genetically improve non-social interaction traits, animals are selected based on their estimated direct breeding value. The direct genetic variance ($\sigma^2_{A_D}$) shows the potential of a non-social interaction trait to respond to selection. As a measure of inheritance, the heritability (h^2) is used, where the direct genetic variance is expressed relative to the phenotypic variance:

$$h^2 = \frac{\sigma^2_{A_D}}{\sigma^2_P}. \quad [1.3]$$

To genetically improve social interaction traits, animals should be selected based on their estimated total breeding value (A_T). The total breeding value combines an animal's direct and indirect breeding value and shows its heritable impact on the mean trait value of a population (Moore *et al.*, 1997; Bijma *et al.*, 2007):

$$A_{T_i} = A_{D_i} + (n-1)A_{I_i}. \quad [1.4]$$

The total genetic variance ($\sigma^2_{A_T}$) shows the potential of a social interaction trait to respond to selection.

$$\sigma^2_{A_T} = \sigma^2_{A_D} + 2(n-1)\sigma_{A_{DI}} + (n-1)^2\sigma^2_{A_I}. \quad [1.5]$$

Specifically, response to selection for social interaction traits (R) depends on the intensity of selection (ι), the accuracy of selection (ρ) and the genetic standard deviation of the total breeding values (σ_{A_T}):

$$R = \iota\rho\sigma_{A_T}. \quad [1.6]$$

Analogous to the classical heritability for non-social interaction traits, the total genetic variance can be expressed relative to the phenotypic variance for social interaction traits (Bergsma *et al.*, 2008):

$$T^2 = \frac{\sigma^2_{A_T}}{\sigma^2_P}. \quad [1.7]$$

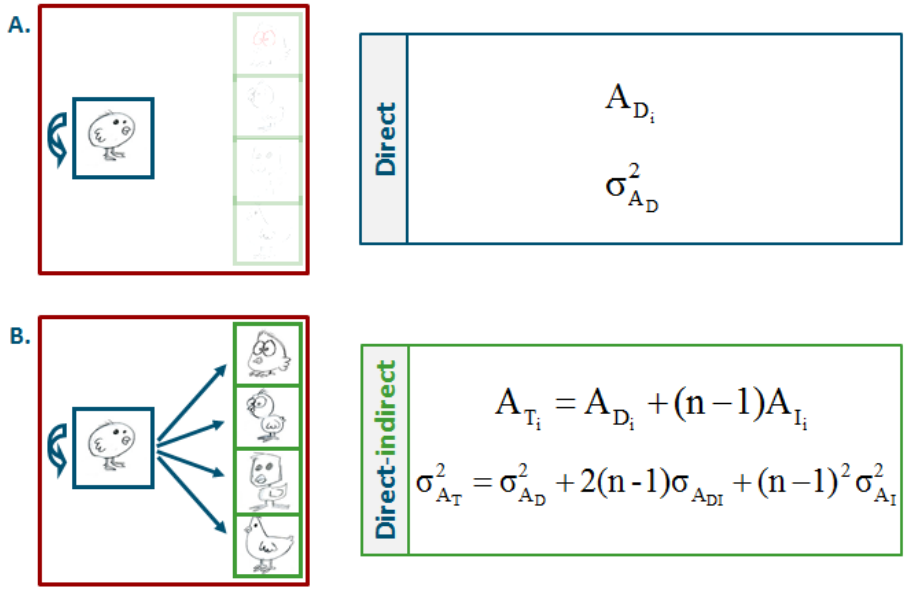

Figure 1.2 Breeding values and genetic variance relevant for response to selection for non-social (A) and social (B) interaction traits

1.3 Feather pecking and cannibalism in laying hens

Feather pecking and cannibalistic behaviour in laying hens impair animal welfare. The behaviour can lead to pulled feathers and tissue damage, both known to be painful to the victim (Savory, 1995). The behaviour also has a negative economic impact, since it results in higher feeding costs, reduced productivity and increased mortality (Leeson and Morrison, 1978; Blokhuis and Wiepkema, 1998; Huber-Eicher and Sebo, 2001). The need to reduce feather pecking and cannibalistic behaviour

increases as more countries ban beak-trimming (Van Horne and Achterbosch, 2008; Nicol *et al.*, 2013). In addition, the problem is more apparent in larger flocks, partly because feather peckers and cannibals have more potential victims to target and partly because the problem is more difficult to constrain (Rodenburg *et al.*, 2004; Fossum *et al.*, 2009). Therefore, the ban on conventional battery cages in the European Union (Council Directive 1999/74/EC) and the shift to alternative housing systems, where animals are kept in larger groups, increases the need for a solution.
Behavioural phenotypes are difficult to record. However, traits like feather condition score or mortality show the consequences of the behaviour and can be recorded more easily. In this thesis, we analysed survival time in non-beak-trimmed laying hens housed in battery cages. This trait is a good example of a social interaction trait, as a bird's chance to survive not only depends on the bird itself, but also depends on the feather pecking and cannibalistic behaviour of its cage mates. Literature shows that there is a genetic component involved in feather pecking and cannibalism, since some genetic lines show more feather pecking and cannibalistic behaviour than others (Blokhuis and Beutler, 1992; Jones *et al.*, 1995; Rodenburg *et al.*, 2010a). In addition, estimated variance components show that both receiving and performing feather pecking and cannibalistic behaviour are heritable. Heritabilities for receiving the behaviour go up to 0.15, while heritabilities for performing the behaviour go up to 0.54 (Cuthbertson, 1980; Bessei, 1984; Kjaer and Sorensen, 1997). Ellen *et al.* (2008) estimated direct and indirect genetic effects for survival time in three purebred White leghorn lines. Laying hens were not beak-trimmed and housed in cages of four. Large and significant indirect genetic effects were found in two out of three lines, *i.e.* W1 and WB. The classical heritability was 0.07 for W1 and 0.10 for WB, while T^2 was 0.19 for W1 and 0.15 for WB. This shows that indirect genetic effects affect the trait and that the potential to respond to selection was underestimated when ignoring IGEs. Finally, selection experiments show that breeding for reduced mortality due to feather pecking and cannibalism is possible (Muir, 1996; Ellen *et al.*, 2013). In conclusion, these studies show that there are good prospects for breeding against mortality due to feather pecking and cannibalistic behaviour.

1.4 Aim and outline of this thesis

This thesis is part of the STW-project entitled "Genetics of social interactions in livestock: improving health, welfare, and productivity in laying hens and pigs". The aim of the project was to understand the inheritance of social interaction traits. This thesis focused on the genetics of social interactions in laying hens. Building on

theoretical work from a previous project, the STW-project focused on the practical implementation. Ideally, the newly obtained knowledge would be implemented in breeding programs, thereby improving welfare and productivity, and reducing mortality in laying hens.

Prior to this thesis, knowledge on direct and indirect genetic effects in laying hens was limited to a study by Ellen *et al.* (2008) who analysed survival time in three purebred lines. However, the ultimate breeding goal is to improve survival in crossbred populations. Therefore, in Chapter 2, the aim was to gain insight in the magnitude of indirect genetic effects for survival time in crossbred laying hens. In Chapter 3, the aim was to quantify the added value of genomic selection for the genetic improvement of survival time in crossbred laying hens.

Survival time in laying hens is recorded on individual birds. However, most production traits in laying hens are recorded on entire groups, resulting in pooled records. It was unclear if indirect genetic effects can be estimated from pooled records. Moreover, it was unclear how individual and pooled records on social interaction traits can be modelled multivariatly. Therefore, in Chapter 4, the aim was to determine whether pooled data can be used to estimate direct, indirect and total genetic effects for social interaction traits. In Chapter 5, the aim was to find a model that yields unbiased genetic parameter estimates for a bivariate analysis of individual and pooled data on social interaction traits.

The general discussion (Chapter 6) addresses two topics. First, I discussed the interpretation of T^2 as a measure of inheritance for social interaction traits. Second, I discussed some of the hurdles we run into when aiming to genetically improve social interaction traits.

2

Indirect genetic effects for survival in domestic chickens (*Gallus gallus*) are magnified in crossbred genotypes and show a parent-of-origin effect

K. Peeters[1], T.T. Eppink[1], E.D. Ellen[1], J. Visscher[2], P. Bijma[1]

[1] Animal Breeding and Genomics Centre, Wageningen University, 6700 AH Wageningen, The Netherlands; [2]Institut de Sélection Animale B.V., Hendrix Genetics, 5830 AC Boxmeer, The Netherlands

Genetics (2012) 192:705-713

Abstract

Through social interactions, individuals can affect one another's phenotype. The heritable effect of an individual on the phenotype of a conspecific is known as an indirect genetic effect (IGE). Although IGEs can have a substantial impact on heritable variation and response to selection, little is known about the genetic architecture of traits affected by IGEs. We studied IGEs for survival time in domestic chickens (*Gallus gallus*), using data on two purebred lines and their reciprocal cross. Birds were kept in groups of four. Feather pecking and cannibalism caused mortality, as beaks were kept intact. Survival time was shorter in crossbreds than in purebreds, indicating outbreeding depression and the presence of non-additive genetic effects. IGEs contributed the majority of heritable variation in crossbreds (87% and 72%) and around half of heritable variation in purebreds (65% and 44%). There was no evidence of dominance variance, neither direct nor indirect. Absence of dominance variance in combination with considerable outbreeding depression suggests that survival time is affected by many loci. Direct-indirect genetic correlations were moderately to highly negative in crossbreds (-0.37 ± 0.17 and -0.83 ± 0.10), but low and not significantly different from zero in purebreds (0.20 ± 0.21 and -0.28 ± 0.18). Consequently, unlike purebreds, crossbreds would fail to respond positively to mass selection. The direct genetic correlation between both crosses was high (0.95 ± 0.23), whereas the indirect genetic correlation was moderate (0.41 ± 0.26). Thus, for IGEs, it mattered which parental line provided the sire and which provided the dam. This indirect parent-of-origin effect appeared to be paternally transmitted and is probably Z-chromosome linked.

Key words: social interactions, indirect genetic effects, survival time, parent-of-origin effect, crossbreeding, laying hens

2.1 Introduction

Social interactions among individuals are widespread in natural and domestic populations (Frank, 2007). When individuals interact, their phenotype can change under the influence of the (behavioural) characteristics of conspecifics. In case these characteristics have a genetic basis, the social environment contains a heritable component (Willham, 1963; Griffing, 1967; Kirkpatrick and Lande, 1989; Moore *et al.*, 1997; Muir, 2005; Bijma *et al.*, 2007). The heritable effect of an individual on the phenotype of a conspecific is known as an indirect genetic effect (IGE) in evolutionary literature, and as an associative, competition or social effect in animal, plant and tree breeding literature (Griffing, 1967; Kirkpatrick and Lande, 1989; Moore *et al.*, 1997; Muir, 2005; Bijma *et al.*, 2007; Van Vleck *et al.*, 2007; Bergsma *et al.*, 2008). The most frequently studied IGE is a maternal genetic effect, which is the heritable environmental effect of a mother on the phenotype of her offspring (Willham, 1963; Cheverud, 1984; Kirkpatrick and Lande, 1989; Koerhuis and Thompson, 1997; Mousseau and Fox, 1998; Eaglen and Bijma, 2009; Bouwman *et al.*, 2010).

The genetic architecture of traits affected by IGEs can differ substantially from ordinary traits. IGEs influence a trait's inheritance and contribute to heritable variation (Hamilton, 1964a,b; Griffing, 1967, 1977; Kirkpatrick and Lande 1989; Moore *et al.*, 1997; Wolf *et al.*, 1998; Bijma and Wade, 2008). Early theoretical work shows that IGEs can explain both positive response to negative selection, *e.g.* evolution of altruism (Hamilton, 1964a,b), and negative response to positive selection, *e.g.* failure of artificial selection for increased or decreased trait values (Griffing, 1967). Those theoretical predictions have been substantiated by selection experiments in animals, plants and bacteria (Wade, 1977; Craig, 1982; Goodnight, 1985; Kyriakou and Fasoulas, 1985; Griffin *et al.*, 2004; Muir, 2005). More recent theoretical work shows that IGEs can contribute substantially to heritable variation, even to the extent that heritable variance exceeds phenotypic variance (Bijma, 2011a). Estimates of indirect genetic variance in beef cattle, pigs and laying hens confirm that IGEs can contribute substantially to heritable variation in agricultural populations (Van Vleck *et al.*, 2007; Bergsma *et al.*, 2008; Chen *et al.*, 2008; Ellen *et al.*, 2008; Chen *et al.*, 2009). These findings are in accordance with predictions of Denison *et al.* (2003), who argued that IGEs are likely to harbour substantial heritable variation, which can be used for genetic improvement. Moreover, the IGE-modelling approach can explain why certain heritable traits, such as success in

pairwise contests, will never respond to selection (Wilson *et al.*, 2011) and allows the quantitative genetic modelling of traits that cannot be attributed to a single individual, such as the number of prey caught by a hunting pack (Bijma, 2011a). The above demonstrates that IGEs can have a big impact on a trait's inheritance and heritable variation. More knowledge of IGEs is needed to predict and understand response to selection in domestic and natural populations.

Traits affected by IGEs can be modelled using either a trait-based approach or a variance component approach. In trait-based models, IGEs are attributed to specific traits and an individual's IGE is the product of its trait values and a coefficient representing the strength of the interaction. These models require knowledge of the social traits that affect the phenotype of a conspecific (Kirkpatrick and Lande, 1989; Moore *et al.*, 1997). In variance component models, in contrast, direct and indirect genetic (co)variances are estimated without knowledge of the social traits that underlie IGEs (Willham, 1963; Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007). A trait-based approach can help us understand the biological mechanism of social interactions. However, when the underlying social traits are unknown or unrecorded, a variance component approach is needed.

At present, knowledge of the magnitude and nature of IGEs is limited (apart from maternal genetic effects; reviewed by Bijma, 2011b). In laying hens, large and statistically significant indirect genetic variances were found for survival time in two out of three investigated purebred lines (Ellen *et al.*, 2008). Results in beef cattle and pigs are diverse. Some studies report large and statistically significant indirect genetic variances, while others report the opposite (Van Vleck *et al.*, 2007; Bergsma *et al.*, 2008; Chen *et al.*, 2008, 2009; Bouwman *et al.*, 2010; Hsu *et al.*, 2010). In addition to additive genetic effects, IGEs might depend on dominance and epistasis, affecting the maintenance of genetic variation and the level of heterosis or inbreeding depression (Lynch and Walsh, 1998). Moreover, IGEs might depend on maternal effects, sex-chromosome linked effects or imprinting, enforcing differences among reciprocal crosses. Further study on the magnitude and nature of IGEs is needed to understand the inheritance of traits affected by IGEs and to optimize genetic improvement in agriculture and aquaculture.

Here we present estimated genetic parameters for survival time in domestic chickens (*Gallus gallus*), using data on two purebred lines and their reciprocal cross. Survival in livestock populations usually has low heritabilities (Dematawega and Berger, 1998; Knol *et al.*, 2002; Quinton *et al.*, 2011). In laying hens, estimates

of genetic parameters for survival during the productive period of group-housed individuals are, to our knowledge, limited to those of Ellen *et al.* (2008). Ellen *et al.* (2008) estimated genetic parameters for survival time in three purebred layer lines. Heritability estimates varied between 0.02 and 0.10. However, Ellen *et al.* (2008) found substantially more heritable variation when accounting for IGEs. To gain knowledge of the nature of IGEs for survival time in domestic chickens, we estimated direct and indirect genetic (co)variances, and investigated whether there is evidence for dominance, epistasis, maternal effects, sex-chromosome linked effects or imprinting.

2.2 Background

This section introduces basic quantitative genetic principles of traits affected by IGEs, using a variance component approach, and introduces the genetic parameters that will be estimated in the next sections. See Table 2.1 for notation.

Classical quantitative genetic theory defines the phenotype (P) as the sum of a genetic (A) and a non-heritable (E) component; $\mathrm{P} = \mathrm{A} + \mathrm{E}$ (Lynch and Walsh, 1998). For traits under the influence of social interactions, the classical model is expanded with IGEs (Willham, 1963; Griffing, 1967). An individual's phenotype now consists of the direct genetic ($\mathrm{A_{D_i}}$) and non-heritable ($\mathrm{E_{D_i}}$) effect of the individual itself (i) and the indirect genetic ($\mathrm{A_{I_j}}$) and non-heritable ($\mathrm{E_{I_j}}$) effects of its group mates (j);

$$\mathrm{P_i} = \mathrm{A_{D_i}} + \mathrm{E_{D_i}} + \sum_{\mathrm{i \neq j}}^{\mathrm{n-1}} \mathrm{A_{I_j}} + \sum_{\mathrm{i \neq j}}^{\mathrm{n-1}} \mathrm{E_{i_j}},$$

where n is the number of group members (Griffing, 1967). With unrelated group members, the phenotypic variance (σ^2_{P}) equals $\sigma^2_{\mathrm{A_D}} + \sigma^2_{\mathrm{E_D}} + (\mathrm{n}-1)\sigma^2_{\mathrm{A_I}} + (\mathrm{n}-1)\sigma^2_{\mathrm{E_I}}$ (Griffing, 1967).

The heritable impact of an individual i on the mean trait value of the population, known as the total breeding value ($\mathrm{A_T}$), consists of its direct breeding value ($\mathrm{A_D}$) and $\mathrm{n}-1$ times its indirect breeding value ($\mathrm{A_I}$);

$\mathrm{A_{T_i}} = \mathrm{A_{D_i}} + (\mathrm{n}-1)\mathrm{A_{I_i}}$.

Consequently, the total heritable variance ($\sigma^2_{\mathrm{A_T}}$), determining a population's potential to respond to selection, equals $\sigma^2_{\mathrm{A_D}} + 2(\mathrm{n}-1)\sigma_{\mathrm{A_{DI}}} + (\mathrm{n}-1)^2\sigma^2_{\mathrm{A_I}}$ (Bijma *et al.*, 2007).

Table 2.1 Notation key

Symbol	Meaning
i - j	Focal individual - Group mates of the focal individual
A_D	Direct genetic effect; Direct breeding value
A_I	Indirect genetic effect; Indirect breeding value
A_T	Total genetic effect; Total breeding value
E_D	Direct non-heritable effect
E_I	Indirect non-heritable effect
σ^2_{Cage}	Cage variance
$\sigma^2_{A_D}$	Direct genetic variance
$\sigma^2_{A_I}$	Indirect genetic variance
$\sigma^2_{A_T}$	Total genetic variance
σ^2_E	Error variance
σ^2_P	Phenotypic variance
$\sigma_{A_{1\_DI}}$; $\sigma_{A_{2\_DI}}$	Direct-indirect genetic covariances within a cross
r_{DI}	Genetic correlation between A_D's and A_I's within a cross
h^2	Heritability
T^2	Total heritable variance relative to phenotypic variance
$\sigma_{A_{1\_D\_2\_I}}$; $\sigma_{A_{2\_D\_1\_I}}$	Direct-indirect genetic covariances between crosses
$\sigma_{A_{12\_D}}$	Direct genetic covariance between crosses
$\sigma_{A_{12\_I}}$	Indirect genetic covariance between crosses
$\sigma_{A_{12\_T}}$	Total genetic covariance between crosses

By analogy of the classical heritability (h^2), $\sigma^2_{A_T}$ can be expressed relative to σ^2_P (Bergsma *et al.*, 2008);
$T^2 = \sigma^2_{A_T}/\sigma^2_P$.

In conclusion, the relevant genetic parameters for traits affected by IGEs are $\sigma^2_{A_D}$, $\sigma_{A_{DI}}$ and $\sigma^2_{A_I}$. Based on these parameters, $\sigma^2_{A_T}$, T^2 and the direct-indirect genetic correlation (r_{DI}) are calculated.

2.3 Materials

Data were provided by the Institut de Sélection Animale B.V., the layer breeding division of Hendrix Genetics. Two commercial purebred White Leghorn layer lines, W1 and WB, were used to produce 15 012 crossbred laying hens of which 7 668 were W1xWB (♂x♀) and 7 344 were WBxW1 (♂x♀). Each cross was produced by randomly mating ~50 sires to ~705 dams, where dams were nested within sires.

Eggs were hatched in two batches. Each batch contained two groups that differed two weeks in age. Post-hatching, chicks were wing-banded, sexed and vaccinated for Infectious Bronchitis and Marek's disease. Their beaks were kept intact. Chicks of the same cross were housed in cages of 60 individuals. At five weeks of age, group size was reduced to 20 individuals. At approximately 17 weeks of age, each batch was placed in a different laying house. The laying houses consisted of four or five double rows. Only eight rows were used per laying house. Consequently, the outer two rows were left empty in one of the laying houses. Each row consisted of three levels, *i.e.* top, middle and bottom. Four hens of the same cross and age were randomly assigned to a cage. A feeding trough was located in front of the cage. Drinking nipples were located in the back of the cage and were shared with back neighbours. Hence, some interaction with back neighbours was possible, but interaction with side neighbours was prevented.

The trait of interest, survival time, was defined as the number of days from the start of the laying period until either death or the end of the experiment, with a maximum of 398 days. Cages were checked daily. Dead hens were removed and the cause of death was determined subjectively. The record was set to missing when the cause of death was clearly unrelated to feather pecking or cannibalism (n=23, birds with broken wings or legs, and birds that were trapped).

In addition, to investigate the impact of crossbreeding on genetic parameters, data on 6 276 W1 and 6 916 WB purebred laying hens, previously analysed by Ellen *et al.* (2008), were reused. More details on the purebred material can be found in Ellen *et al.* (2008).

Table 2.2 Number of laying hens (n), their survival (%) with standard error, and their average number of survival days with standard error for each cross, level and row within each laying house

	Laying house 1			Laying house 2		
	n	Survival	Survival days	n	Survival	Survival days
Total	8 072	50 ± 1	273 ± 2	6 940	62 ± 1	307 ± 2
Cross						
W1xWB	4 292	53 ± 1	290 ± 2	3 376	68 ± 1	328 ± 2
WBxW1	3 780	46 ± 1	254 ± 3	3 564	56 ± 1	287 ± 2
Level						
Top	2 444	42 ± 1	244 ± 3	356	72 ± 2	338 ± 6
Middle	2 596	51 ± 1	279 ± 3	3 280	62 ± 1	306 ± 2
Bottom	3 032	55 ± 1	290 ± 3	3 304	62 ± 1	305 ± 2
Row						
1	176	55 ± 4	282 ± 11	632	65 ± 2	314 ± 5
2	1 224	51 ± 1	269 ± 4	884	60 ± 2	305 ± 4
3	1 220	50 ± 1	276 ± 4	888	65 ± 2	314 ± 4
4	1 220	51 ± 1	278 ± 4	888	59 ± 2	297 ± 5
5	1 224	58 ± 1	305 ± 4	888	61 ± 2	305 ± 5
6	1 224	49 ± 1	277 ± 4	884	56 ± 2	285 ± 5
7	1 224	44 ± 1	249 ± 4	884	57 ± 2	297 ± 5
8	560	41 ± 2	231 ± 7	992	75 ± 1	339 ± 4

2.4 Methods

A linear mixed model was used to estimate genetic parameters for survival time (motivated in the *Results and Discussion*). To determine which fixed effects to include in the model, a general linear model was run in SAS v9.1 (SAS Institute Inc., 2003). First, an interaction term for each laying house by row by level combination was included to correct for infrastructural effects (*e.g.* differences in light intensity). Second, a fixed effect for the content of the back cage was included, which was either empty or contained hens. Third, a covariate for the average number of survival days in the back cage was included. The model was then extended with random effects in ASReml v3.0 (Gilmour *et al.*, 2009).

To investigate whether genetic parameters differ between both crosses, a bivariate animal model was used in which survival time was analysed as a statistically different trait for each cross.

2.4.1 Direct animal model

The following model was used to estimate direct genetic parameters:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & 0\\ 0 & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{1\_D} & 0\\ 0 & \mathbf{Z}_{2\_D}\end{bmatrix}\begin{bmatrix}\mathbf{a}_{1\_D}\\ \mathbf{a}_{2\_D}\end{bmatrix} + \begin{bmatrix}\mathbf{V}_1 & 0\\ 0 & \mathbf{V}_2\end{bmatrix}\begin{bmatrix}\mathbf{cage}_1\\ \mathbf{cage}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix},$$

where subscript 1 refers to W1xWB and subscript 2 refers to WBxW1; $\mathbf{y}$ is a vector of observations; $\mathbf{X}$ is an incidence matrix linking observations to fixed effects; $\mathbf{b}$ is a vector of fixed effects; $\mathbf{Z}_D$ is an incidence matrix linking an animal's phenotype to its own A_D; $\mathbf{a}_D$ is a vector of A_D's; $\mathbf{V}$ is an incidence matrix linking observations to random cage effects; $\mathbf{cage}$ is a vector of independent random cage effects; and $\mathbf{e}$ is a vector of residuals.

The direct genetic covariance structure was:

$$\mathrm{Var}\begin{bmatrix}\mathbf{a}_{1\_D}\\ \mathbf{a}_{2\_D}\end{bmatrix} = \begin{bmatrix}\sigma^2_{A_{1\_D}} & \sigma_{A_{12\_D}}\\ \sigma_{A_{12\_D}} & \sigma^2_{A_{2\_D}}\end{bmatrix} \otimes \mathbf{A},$$

where $\sigma^2_{A_{1\_D}}$ is the direct genetic variance for W1xWB; $\sigma^2_{A_{2\_D}}$ is the direct genetic variance for WBxW1; $\sigma_{A_{12\_D}}$ is the direct genetic covariance between crosses; and $\mathbf{A}$ is the matrix of additive genetic relationships between individuals, based on five generations of pedigree.

2.4.2 Direct-indirect animal model

The following model was used to estimate direct and indirect genetic parameters:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & 0\\ 0 & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{1\_D} & 0\\ 0 & \mathbf{Z}_{2\_D}\end{bmatrix}\begin{bmatrix}\mathbf{a}_{1\_D}\\ \mathbf{a}_{2\_D}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{1\_I} & 0\\ 0 & \mathbf{Z}_{2\_I}\end{bmatrix}\begin{bmatrix}\mathbf{a}_{1\_I}\\ \mathbf{a}_{2\_I}\end{bmatrix} + \begin{bmatrix}\mathbf{V}_1 & 0\\ 0 & \mathbf{V}_2\end{bmatrix}\begin{bmatrix}\mathbf{cage}_1\\ \mathbf{cage}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix},$$

where the vectors and incidence matrices correspond to those in the direct animal model; $\mathbf{Z}_I$ is an incidence matrix linking an animal's phenotype to its cage mates' A_I; and $\mathbf{a}_I$ is the vector of A_I's.

The direct-indirect genetic covariance structure was:

$$\mathrm{Var}\begin{bmatrix}\mathbf{a}_{1\_D}\\ \mathbf{a}_{2\_D}\\ \mathbf{a}_{1\_I}\\ \mathbf{a}_{2\_I}\end{bmatrix} = \begin{bmatrix}\sigma^2_{A_{1\_D}} & \sigma_{A_{12\_D}} & \sigma_{A_{1\_DI}} & \sigma_{A_{1\_D\_2\_I}}\\ \sigma_{A_{12\_D}} & \sigma^2_{A_{2\_D}} & \sigma_{A_{2\_D\_1\_I}} & \sigma_{A_{2\_DI}}\\ \sigma_{A_{1\_DI}} & \sigma_{A_{2\_D\_1\_I}} & \sigma^2_{A_{1\_I}} & \sigma_{A_{12\_I}}\\ \sigma_{A_{1\_D\_2\_I}} & \sigma_{A_{2\_DI}} & \sigma_{A_{12\_I}} & \sigma^2_{A_{2\_I}}\end{bmatrix} \otimes \mathbf{A},$$

where $\sigma^2_{A_{1\_I}}$ is the indirect genetic variance for W1xWB; $\sigma^2_{A_{2\_I}}$ is the indirect genetic variance for WBxW1; $\sigma_{A_{12\_I}}$ is the indirect genetic covariance between crosses; $\sigma_{A_{1\_DI}}$ is the direct-indirect genetic covariance within W1xWB; $\sigma_{A_{2\_DI}}$ is the direct-indirect genetic covariance within WBxW1; $\sigma_{A_{1\_D\_2\_I}}$ is the genetic covariance between the direct genetic effect of W1xWB and the indirect genetic effect of WBxW1; and $\sigma_{A_{2\_D\_1\_I}}$ is the genetic covariance between the direct genetic effect of WBxW1 and the indirect genetic effect of W1xWB.

The above direct-indirect animal model was also used univariate to estimate genetic parameters in the purebred parental lines, *i.e.* W1 and WB. The purebred data were previously analysed by Ellen *et al.* (2008), but with a different experimental time span and a slightly different model. Therefore, the purebred data were reanalysed using the same experimental time span (398 instead of 447 days) and the above direct-indirect animal model.

2.4.3 Direct-indirect animal model with non-genetic maternal effects

The following model was used to estimate non-genetic direct and indirect maternal effects:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & 0\\ 0 & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{1\_D} & 0\\ 0 & \mathbf{Z}_{2\_D}\end{bmatrix}\begin{bmatrix}\mathbf{a}_{1\_D}\\ \mathbf{a}_{2\_D}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{1\_I} & 0\\ 0 & \mathbf{Z}_{2\_I}\end{bmatrix}\begin{bmatrix}\mathbf{a}_{1\_I}\\ \mathbf{a}_{2\_I}\end{bmatrix} + \begin{bmatrix}\mathbf{W}_{1\_D} & 0\\ 0 & \mathbf{W}_{2\_D}\end{bmatrix}\begin{bmatrix}\mathbf{dam}_{1\_D}\\ \mathbf{dam}_{2\_D}\end{bmatrix} + \begin{bmatrix}\mathbf{W}_{1\_I} & 0\\ 0 & \mathbf{W}_{2\_I}\end{bmatrix}\begin{bmatrix}\mathbf{dam}_{1\_I}\\ \mathbf{dam}_{2\_I}\end{bmatrix} + \begin{bmatrix}\mathbf{V}_1 & 0\\ 0 & \mathbf{V}_2\end{bmatrix}\begin{bmatrix}\mathbf{cage}_1\\ \mathbf{cage}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix},$$

where the vectors and incidence matrices correspond to those in the direct-indirect animal model; $\mathbf{W}_D$ is an incidence matrix linking an animal's phenotype to its own dam; $\mathbf{dam}_D$ is the vector of independent random direct maternal effects; $\mathbf{W}_I$ is an incidence matrix linking an animal's phenotype to its cage mates' dam; and $\mathbf{dam}_I$ is the vector of independent random indirect maternal effects.

Non-genetic maternal effects account for the covariance between maternal siblings apart from their additive genetic relationship. Such a covariance may arise because of shared maternal environment, causing full-sibs to express similar direct or indirect effects. Moreover, because of the nested mating structure, the maternal effect also accounts for the covariance among siblings due to non-additive effects such as dominance and epistasis, both direct and indirect. Omitting non-genetic maternal effects from the model may cause overestimation of the additive genetic variance.

2.5 Results and discussion

2.5.1 Descriptive statistics

A significant difference in survival was found between both crosses. Up to day 398, 61% of the W1xWB hens survived, while only 51% of the WBxW1 hens survived (Figure 2.1). A significant difference in survival was also observed between laying houses. In laying house one, 50% of the hens survived, while in laying house two, 62% of the hens survived (Table 2.2). This is probably related to the higher light intensity in laying house one (Ellen *et al.*, 2008), which is known to evoke feather pecking and cannibalism (Hughes and Duncan, 1972; Savory, 1995; Kjaer and Vestergaard, 1999).

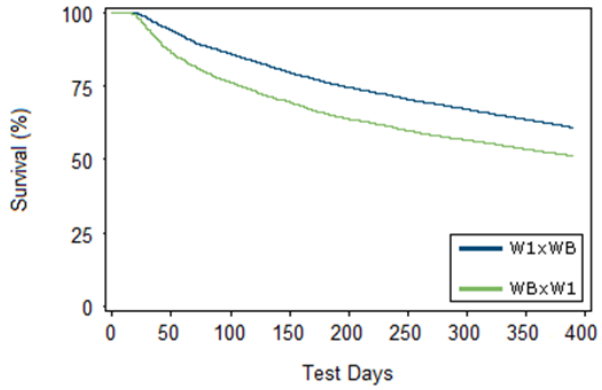

Figure 2.1 The Kaplan-Meier survival curve; plotting survival (%) against the number of test days (0-398) for the reciprocal cross of two purebred layer lines

2.5.2 Model

When analysing survival time, one should ideally use survival analysis methodology to account for the skewed and censored nature of the data (Kalbfleisch and Prentice, 2002). However, survival analysis software does not yet allow the inclusion of both direct and indirect genetic effects. To circumvent this problem, Ellen *et al.* (2010) proposed a two-step procedure, which combines survival analysis (step one) with a linear mixed model (step two). In the first step, only a direct genetic effect is modelled using survival analysis. The estimate of this effect is then used to create a pseudo-record. In the second step, the pseudo-record is modelled using a linear direct-indirect mixed model. Cross validation results showed that the ordinary linear mixed model had the same predictive ability of breeding values as the two-step procedure (Ellen *et al.*, 2010). Ellen *et al.* (2010) did not investigate variance component estimation with the two-step procedure. When analysing the current data with the two-step procedure, the estimated genetic parameters showed to be highly dependent on the animal to which the phenotype was allocated in the first step, *i.e.* the focal animal or one of its group mates. The decision whether to fit a direct or an indirect genetic effect in the first step should not affect the final outcome. Hence, more research is needed to optimize the two-step procedure for variance component estimations. Therefore, the decision was made to fit a linear mixed model.

2.5.3 Genetic parameters within crosses

Table 2.3 shows the estimated parameters, for crossbreds, from the direct animal model. The additive direct genetic variance was highly significant in both crosses. The heritability was 0.05 for W1xWB and 0.06 for WBxW1. When comparing both crosses, all variance components were smaller in W1xWB. This is a direct consequence of the difference in mean survival between both crosses. Since W1xWB has a higher survival, more observations were censored and less variation was observed.

Table 2.3 also shows the estimated parameters from the direct-indirect animal model. The model fitted the data considerably better than the direct animal model (χ-square test; P<0.001). Both $\sigma^2_{A_D}$ and $\sigma^2_{A_I}$ were highly significant in both crosses. For the same reason as for the direct animal model, all variance components were smaller in W1xWB. Although $\sigma^2_{A_I}$ has a similar magnitude as $\sigma^2_{A_D}$, its contribution to σ^2_P is three times larger and its contribution to $\sigma^2_{A_T}$ is nine times larger ($n-1=3$ and $(n-1)^2=9$, see *Background*). T^2 was 0.26 for W1xWB and 0.17 for WBxW1. These values substantially exceed the ordinary (direct) h^2, indicating that the majority of heritable variation was hidden in the direct animal model. The difference in T^2 between both crosses is primarily due to the difference in $\sigma_{A_{DI}}$, rather than the difference in direct or indirect genetic, and non-heritable variance. The direct-indirect genetic correlation was moderately negative for W1xWB (-0.37), but highly negative for WBxW1 (-0.83). These negative correlations indicate that individuals with a positive (direct) breeding value for their own survival have, on average, a negative (indirect) breeding value for the survival of their cage mates and *vice versa*. This can be interpreted as heritable competition. With heritable competition, selection for direct genetic effects results in a negative indirect genetic response and potentially in a negative net response in the phenotype. When unrelated group members are selected based on their own performance, the realized heritability equals $\left[\sigma^2_{A_D}+(n-1)\sigma_{A_{DI}}\right]/\sigma^2_P$ (Griffing, 1967). Based on the estimates in Table 2.3, mass selection would result in a realized heritability of 0.00 for W1xWB and -0.06 for WBxW1. Hence, despite substantial heritable variance, W1xWB would fail to respond to mass selection and WBxW1 would respond in the opposite direction. The large difference between the realized heritability for mass selection and T^2 demonstrates that breeders need to adapt their selection criterion in order to achieve positive response to selection in these crosses.

Table 2.3 Estimated parameters with standard error, for crossbreds, from a direct and a direct-indirect animal model

	Direct animal model		Direct-indirect animal model	
	W1xWB	WBxW1	W1xWB	WBxW1
σ^2_{Cage}	2 764 ± 205	2 914 ± 265	1 984 ± 260	2 379 ± 306
$\sigma^2_{A_D}$	711 ± 179	1 292 ± 281	536 ± 152	997 ± 226
$\sigma_{A_{DI}}$			-197 ± 93	-726 ± 140
$\sigma^2_{A_I}$			536 ± 109	767 ± 148
σ^2_E	12 326 ± 270	17 109 ± 393	11 732 ± 298	15 655 ± 460
σ^2_P	15 802 ± 273	21 315 ± 373	15 860 ± 289	21 332 ± 401
h^2; T^2	0.05 ± 0.01	0.06 ± 0.01	0.26 ± 0.06	0.17 ± 0.05
r_{DI}			-0.37 ± 0.17	-0.83 ± 0.10
$\sigma_{A_{12\_D}}$	770 ± 253		697 ± 209	
$\sigma_{A_{12\_I}}$			261 ± 172	
$\sigma_{A_{1\_D\_2\_I}}$			-285 ± 185	
$\sigma_{A_{2\_D\_1\_I}}$			90 ± 190	
r_{12_D}	0.80 ± 0.22		0.95 ± 0.23	
r_{12_I}			0.41 ± 0.26	
r_{12_T}			0.64 ± 0.31	

Estimates of the non-genetic direct and indirect maternal effects were small and statistically non-significant. This implies that common environmental effects due to the dam are negligible. Moreover, variance due to dominance and epistasis seem negligible.

2.5.4 Purebred-crossbred comparison

In crossbreds, 61% of the W1xWB and 51% of the WBxW1 hens survived up to day 398. In the purebred lines, survival up to day 398 was 64% for W1 and 58% for WB. Thus, on average, survival time was shorter in crossbreds than in purebreds. Because test circumstances were similar for pure- and crossbreds (same stables, different year), the decrease in survival is most probably due to non-additive genetic effects, rather than environmental effects. Non-additive genetic effects and negligible non-additive genetic variance seem to contradict each other. However, if many loci influence the trait, dominance variance can be small, despite substantial (negative) heterosis. Because heterosis is proportional to the dominance effect (d) at a locus ($H \sim \sum_{Loci} d$) and dominance variance is proportional to d^2 ($\sigma_D^2 \sim \sum_{Loci} d^2$) dominance variance will decrease when the number loci increases and heterosis is constant (Robertson *et al.*, 1983; Falconer and Mackay, 1996). Hence, this suggests that survival time is affected by many loci, which is consistent with results from Biscarini *et al.* (2010a) who found 11 direct QTL and 81 indirect QTL for feather condition score, which is a precursor of survival.

Table 2.4 shows the estimated parameters, for purebreds, from the direct-indirect animal model. T^2 was 0.19 for W1 and 0.16 for WB. These values are slightly lower than in crossbreds. The underlying parameters, however, showed substantial differences. Although $\sigma^2_{\mathrm{A_D}}$was similar in pure- and crossbreds, $\sigma^2_{\mathrm{A_I}}$ was two to six times larger in crossbreds than in purebreds. Moreover, $\mathrm{r_{DI}}$ was low and not significantly different from zero in purebreds, but moderately to highly negative in crossbreds. Although T^2 was similar in pure- and crossbreds, the contribution of IGEs to $\sigma^2_{\mathrm{A_T}}$ differed. The contribution of $2(\mathrm{n}-1)\sigma_{\mathrm{A_{DI}}} + (\mathrm{n}-1)^2\sigma^2_{\mathrm{A_I}}$ to $\sigma^2_{\mathrm{A_T}}$ was 87% in W1xWB and 72% in WBxW1, while it was 65% in W1 and 44% in WB. Moreover, the realized heritability in case of mass selection differed. The realized heritability was 0.00 for W1xWB and -0.06 for WBxW1, while it was 0.08 for W1 and 0.06 for WB. This indicates that IGEs in crossbreds contribute more to heritable variation and have more impact on response to selection than in the parental purebred lines.

Table 2.4 Estimated parameters with standard error, for purebreds, from a direct-indirect animal model

	W1	WB
σ^2_{Cage}	803 ± 161	1 200 ± 237
$\sigma^2_{A_D}$	656 ± 161	1 400 ± 299
$\sigma_{A_{DI}}$	51 ± 58	-161 ± 104
$\sigma^2_{A_I}$	100 ± 39	228 ± 71
σ^2_E	7 976 ± 205	12 686 ± 364
σ^2_P	9 735 ± 187	15 971 ± 297
T^2	0.19 ± 0.06	0.16 ± 0.05
r_{DI}	0.20 ± 0.21	-0.28 ± 0.18

Unfortunately, because of a lack of close pedigree links between pure- and crossbreds, purebred-crossbred correlations could not be estimated. The large difference in r_{DI} between pure- and crossbreds implies that the purebred-crossbred correlation must be smaller than one, at least for one of the effects (*i.e.* direct or indirect). This indicates the presence of non-additive effects such as dominance or epistasis (Wei *et al.*, 1991), or parent-of-origin effects such as sex-chromosome linked effects or imprinting.

2.5.5 Genetic parameters between crosses

The genetic correlation between the A_T's of both crosses (r_{12_T}; see *Appendix* for derivation) was moderate (0.64) and not significantly different from one (Table 2.3). Underlying, the genetic correlation between direct genetic effects (r_{12_D}) was high (0.95) and not significantly different from one, while the genetic correlation between indirect genetic effects (r_{12_I}) was moderate (0.41) and significantly different from one (Table 2.3). This moderate genetic correlation indicates that, for IGEs, it mattered which parental line provided the sire and which provided the dam, *i.e.* an indirect parent-of-origin effect.

So far, parent-of-origin effects have been reported for direct effects only. In chickens, direct parent-of-origin effects have been found for feed intake, body weight, sexual maturity, egg production traits, egg quality traits and viability (Fairfull *et al.*, 1983; Fairfull and Gowe, 1986; Ledur *et al.*, 2002; Tuiskula-Haavisto *et al.*, 2004). Parent-of-origin effects can have multiple underlying causes, such as (cytoplasmatic) maternal effects, sex-chromosome linked effects or imprinting

(Fairfull *et al.*, 1983; Fairfull and Gowe, 1986; Tuiskula-Haavisto *et al.*, 2004). Because maternal variances were small and statistically non-significant, they can be excluded as a potential cause of the indirect parent-of-origin effect found here.
Comparing pure- and crossbred data revealed that the cross with the highest survival (W1xWB) received the paternal chromosome from the pure line with the highest survival (W1) and *vice versa* (Figure 2.2). This result suggests that part of the genes affecting survival time is located on the paternal sex-chromosome (the Z-chromosome, which carries more genetic information than the W-chromosome) or is maternally imprinted. This result agrees with findings of Rodenburg *et al.* (2003), who reported a higher sire-based than dam-based heritability for feather pecking. Severe feather pecking can kill the recipient (Savory, 1995) and has a major impact on survival time.

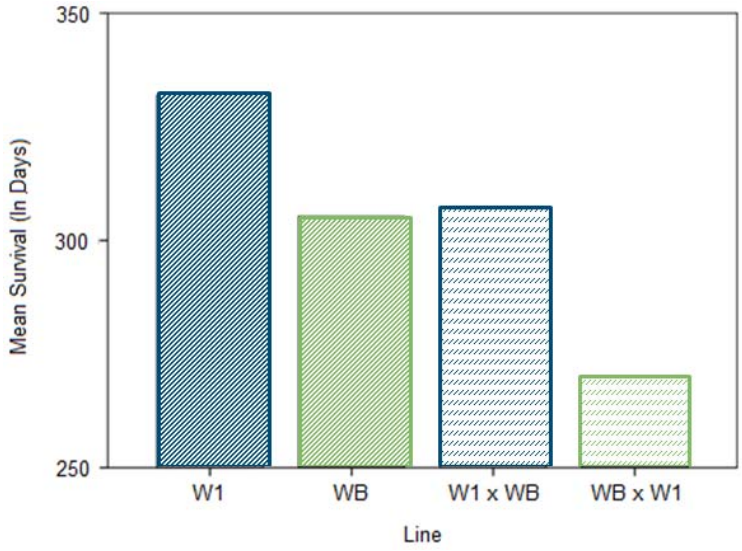

Figure 2.2 Mean survival (in days) for the purebred lines W1 and WB, and their reciprocal cross W1xWB and WBxW1

Sex-chromosomes are known to have a substantial impact on sex-specific behavioural characteristics (Xu *et al.*, 2002; Gatewood *et al.*, 2006). This could also apply for feather pecking and cannibalism, which is more common in females than in males (Hughes, 1973; Jensen *et al.*, 2005). On the one hand, sex-chromosomes contain genes that regulate the expression of gonadal steroid hormones. Hughes (1973) observed that the simultaneous admission of oestrogen and progesterone resulted in more feather pecking and cannibalism, while the admission of

testosterone had the opposite effect. On the other hand, sex-chromosomes contain genes that are not involved in male or female determination, but do affect sex-specific characteristics (Gatewood *et al.*, 2006). These genes can reinforce differences between males and females as, despite a certain degree of dosage compensation, certain parts of the chromosome remain unequally expressed in males and females (Xu *et al.*, 2002; Arnold *et al.*, 2008). Biscarini *et al.* (2010a) found evidence for Z-chromosome linked IGEs in an association study on feather condition score in laying hens. Feather condition score serves as a measure for the severity of feather pecking. Biscarini *et al.* (2010a) identified 81 QTL for IGEs, of which six were located on the Z-chromosome. Once more, this suggests that IGEs for survival time are Z-chromosome linked. On the basis of these observations, the decision was made to perform a sex-chromosome linked analysis. However, this model failed to converge.

Alternatively, maternal imprinting, where only paternally inherited alleles are expressed, could explain the observed parent-of-origin effect. Imprinting in animals is assumed to be a phenomenon exclusive for placental-marsupial mammals, fish and insects, expressed at the embryonic or postnatal stage (Reik and Walter, 2001). However, there are indications that imprinting occurs in birds as well (Reik and Walter, 2001; Tuiskula-Haavisto *et al.*, 2004; Tuiskula-Haavisto and Vilkki, 2007; Úbeda and Gardner, 2010). Moreover, imprinting is recently linked to social behaviour in later stages of life (Garfied *et al.*, 2011). But, because of a lack of biological evidence, through expression studies at RNA or protein level, imprinting is an unlikely explanation for the indirect parent-of-origin effect observed here.

If IGEs for survival time are indeed Z-chromosome linked, this could cause the sire variance to exceed the dam variance. This would occur only if the causal genes on the Z-chromosome are still segregating within the pure lines. When both pure lines carry different IGEs on the Z-chromosome, but those do not segregate within pure lines, then the sire and dam variance would be equal, but r_{12_I} may still be smaller than one. To investigate this issue, $\mathbf{a}_I$ in the above direct-indirect animal model was replaced by an indirect genetic sire and dam effect (model not shown). Results showed no consistent or significant difference between sire and dam variance. This suggests that IGE genes on the Z-chromosome do not segregate within the pure lines. Alternatively, this issue could be explained by a combination of dominance variance and IGEs segregating on the Z-chromosome within pure lines. In theory, dominance could inflate the dam variance, while IGEs on the Z-chromosome could inflate the sire variance by approximately the same amount, resulting in a similar sire and dam variance. Genome-wide association studies are needed to further investigate the genetic architecture of survival time in chickens.

2.6 Acknowledgements

This research was financially supported by the Netherlands Organization for Scientific Research (NWO) and coordinated by the Dutch Technology Foundation (STW). The Institut de Sélection Animale B.V., a Hendrix Genetics Company, provided the data and was closely involved in this research.
The authors thank Abe Huisman, Addie Vereijken, Bas Rodenburg, Duur Aanen, Ewa Sell-Kubiak, Johan van Arendonk, Naomi Duijvesteijn and Sophie Eaglen for their valuable input.

2.7 Appendix

To quantify the similarity between crosses, the genetic correlation between A_T's (r_{12_T}) needed to be calculated. r_{12_T} is dependent on the total heritable variance within crosses ($\sigma^2_{A_{1_T}}$and $\sigma^2_{A_{2\_T}}$) and the total genetic covariance between crosses ($\sigma_{A_{12\_T}}$) , as $r_{12\_T} = \sigma_{A_{12\_T}} / \sqrt{\sigma^2_{A_{1\_T}} \sigma^2_{A_{2\_T}}}$.

With $\sigma^2_{A_{1\_T}} = \mathrm{Var}(A_{1\_D} + (n-1)A_{1\_I})$, it follows that $\sigma^2_{A_{1\_T}} = \sigma^2_{A_{1\_D}} + 2(n-1)\sigma_{A_{1\_DI}} + (n-1)^2\sigma^2_{A_{1\_I}}$. And with $\sigma^2_{A_{2\_T}} = \mathrm{Var}(A_{2\_D} + (n-1)A_{2\_I})$, it follows that $\sigma^2_{A_{2\_T}} = \sigma^2_{A_{2\_D}} + 2(n-1)\sigma_{A_{2\_DI}} + (n-1)^2\sigma^2_{A_{2\_I}}$.

With $\sigma_{A_{12\_T}} = \mathrm{Covar}(A_{1\_D} + (n-1)A_{1\_I}; A_{2\_D} + (n-1)A_{2\_I})$, it follows that $\sigma_{A_{12\_T}} = \sigma_{A_{12\_D}} + (n-1)\sigma_{A_{1\_D\_2\_I}} + (n-1)\sigma_{A_{2\_D\_1\_I}} + (n-1)^2\sigma_{A_{12\_I}}$.

Therefore, $r_{12\_T} = \dfrac{\sigma_{A_{12\_D}} + (n-1)\sigma_{A_{1\_D\_2\_I}} + (n-1)\sigma_{A_{2\_D\_1\_I}} + (n-1)^2\sigma_{A_{12\_I}}}{\sqrt{\left(\sigma^2_{A_{1\_D}} + 2(n-1)\sigma_{A_{1\_DI}} + (n-1)^2\sigma^2_{A_{1\_I}}\right)\left(\sigma^2_{A_{2\_D}} + 2(n-1)\sigma_{A_{2\_DI}} + (n-1)^2\sigma^2_{A_{2\_I}}\right)}}$.

3

Single-step GBLUP

for survival time in crossbred laying hens

K. Peeters[1,2], B.O. Fragomeni[3], P. Bijma[1], W.M. Muir[4], I. Misztal[3]

[1] Animal Breeding and Genomics Centre, Wageningen University, 6700 AH Wageningen, The Netherlands; [2] Research and Technology Centre, Hendrix Genetics, 5831 CK Boxmeer, The Netherlands; [3] Department of Animal and Dairy Science, University of Georgia, Athens GA 30602, United States of America; [4] Department of Animal Science, Purdue University, West Lafayette IN 47906, United States of America

To be submitted

Abstract

Mortality due to feather pecking and cannibalistic behaviour is an important economic and welfare problem in non-beak-trimmed laying hens. Whether or not a bird dies depends on the direct (genetic) effect of the animal itself and the indirect (genetic) effect of its group mates. To improve the trait, animals should be selected based on their total breeding value, which is a combination of their direct and indirect breeding value. In this study, we estimated genetic parameters for survival time using 50 590 records on the crossbred offspring of three sire lines. Genotypic data (60k) were available on the sires. A sire-line specific analysis was run. Either the complete dataset censored at 347 days or a subset censored at 414 days was used. We quantified the impact of the moment of censoring on T^2, which is the total heritable variance relative to the phenotypic variance. Furthermore, we calculated the change in accuracy obtained by using genotypic information in addition to pedigree information, and by delaying the moment of censoring. T^2 for survival time varied from 0.12 to 0.30. Genetic improvement can, therefore, help to decrease mortality due to feather pecking and cannibalism. T^2 was not affected by the moment of censoring. By using genotypic information in addition to pedigree information, the accuracy of the estimated total breeding values increased by 20 up to 110% for sires without progeny information, showing the added value of genomic selection. Moreover, it could be beneficial to extend the moment of censoring to create more phenotypic variation. However, if this implies deleting large batches of data with a shorter experimental time span, the benefit could be lost due to loss of power.

Key words: social interactions, survival time, laying hens, censoring, ssGBLUP

3.1 Introduction

Mortality due to feather pecking and cannibalistic behaviour in laying hens is an economic and animal welfare problem that becomes increasingly important as more countries ban beak-trimming (Leeson and Morrison, 1978; Blokhuis and Wiepkema, 1998; Huber-Eicher and Sebö, 2001; Van Horne and Achterbosch, 2008; Nicol *et al.*, 2013). This unwanted behaviour among laying hens causes feather loss, tissue damage and even death. The problem is more common in alternative housing systems, such as floor housing and aviaries, than in conventional battery-cages (Rodenburg *et al.*, 2004; Fossum *et al.*, 2009). Therefore, the ban on conventional battery-cages in the European Union (Council Directive 1999/74/EC) makes the problem more apparent.

Mortality due to feather pecking and cannibalism is multifactorial. Diet, group size, light regime and other environmental effects are known to influence the trait (Keeling and Jensen, 1995; Van Krimpen *et al.*, 2005; Mohammed *et al.*, 2010; Rodenburg *et al.*, 2004, 2013; Nicol *et al.*, 2013). Also the genetic background of the laying hens influences the trait, as some genetic lines show more feather pecking and cannibalistic behaviour than others (Blokhuis and Beutler, 1992; Jones *et al.*, 1995; Rodenburg *et al.*, 2010a). Selection experiments and variance component estimations show that there are good prospects for genetic improvement (Craig and Muir, 1996; Muir, 1996; Ellen *et al.*, 2008, 2013, 2014; Peeters *et al.*, 2012, Alemu *et al.*, submitted). Mortality due to feather pecking and cannibalism is a social interaction trait (Ellen *et al.*, 2008; Peeters *et al.*, 2012), where a bird's chance to survive not only depends on the tendency of the bird to be a victim, *i.e.* the direct (genetic) effect, but also depends on the tendency of its cage mates to be aggressors, *i.e.* the indirect (genetic) effect (Griffing, 1967). To improve the trait, animals should be selected based on their total breeding value, which is a combination of an animal's direct and indirect breeding value (Bijma *et al.*, 2007). The expected response to selection may increase drastically when social interactions are taken into account. In the literature, the detected heritable variation for survival time in non-beak-trimmed group-housed laying hens increased up to 400% when taking social interactions into account (Ellen *et al.*, 2008; Peeters *et al.*, 2012).

Several studies have estimated the total genetic variance for survival time in non-beak-trimmed group-housed laying hens (Ellen *et al.*, 2008; Peeters *et al.*, 2012, Alemu *et al.*, submitted). Analysis of survival time suffers from censoring, *i.e.* the phenomenon where some individuals are still alive at the end of the experiment and will receive the same phenotype. The moment of censoring differed between

studies and ranged from 372 up to 447 days. Ellen *et al.* (2008) analysed survival time censored at 447 days in three purebred lines. The total genetic variance relative to the phenotypic variance (T^2) varied from 0.06 to 0.19. Peeters *et al.* (2012) analysed survival time censored at 398 days in a reciprocal cross. Depending on the cross, T^2 was 0.17 or 0.26. Alemu *et al.* (submitted) analysed survival time censored at 372 days in the crossbred offspring of two sire lines. Depending on the sire line, T^2 was 0.18 or 0.22. In all three studies, the moment of censoring was chosen based on the batch (*i.e.* group of animals that was hatched together and were simultaneously put in the laying house) with the shortest experimental timespan. Delaying the moment of censoring would be beneficial because the phenotypic variance increases and fewer animals are censored. However, delaying the moment of censoring also implies that batches with a short experimental timespan will be deleted. The impact of delaying the moment of censoring on T^2 and the accuracy of breeding values has not yet been studied.

Survival time in non-beak-trimmed group-housed laying hens shows good prospects for genomic selection. Pure line males and females are usually kept in individual cages. Their crossbred female offspring, however, are kept in group cages. Progeny-testing implies longer generation intervals, as results will not be available until the pure line animals are +/- 700 days old (assuming an experimental time span of 400 days). Hence, genomic selection could allow us to decrease the generation interval while maintaining a certain level of accuracy, potentially increasing response to selection (Meuwissen *et al.*, 2001; Schaeffer, 2006). In brown layers, Alemu *et al.* (submitted) showed that the accuracy of the estimated breeding values for survival time increased up to 40% by using genotypic information in addition to pedigree information. In white layers, the potential increase in accuracy of the estimated breeding values for survival time has not yet been studied.

Here we estimate genetic parameters for survival time in the crossbred offspring of three sire lines (W1, W5 and W6) using single-step genomic BLUP (Misztal *et al.*, 2009; Legarra *et al.*, 2009; Christensen and Lund, 2009). We quantify the impact of the moment of censoring on the estimated variance components and T^2. Furthermore we calculate the change in accuracy when adding genotypic information in addition to pedigree information and by delaying the moment of censoring.

3.2 Materials

Data were provided by the Institut de Sélection Animale B.V., the layer breeding division of Hendrix Genetics. Phenotypic data were available on 50 590 White Leghorn laying hens (*Gallus gallus*) from 13 different crosses (Table 3.1). Pedigree information was available on the sire side only and traced back six generations. Genotypic data (60k) were available on the sires only. Out of the 1 335 sires used, 1 144 were genotyped. Crosses were clustered by sire line (W1, W5 and W6) to allow sire line specific analyses.

Table 3.1 Number of phenotypic records per cross and per sire line

		Sire lines		
		W1	W5	W6
Dam lines	WA	4 155	2 685	2 205
	WB	5 610	1 345	1 665
	WC	5 035	-	1 735
	WD	5 720	-	-
	WG	-	4 175	5 460
	WH	-	5 290	5 510
	Total	20 520	13 495	16 575

Eggs were hatched in eight batches. The first batch was hatched in 2007, the last batch was hatched in 2011. Post-hatching, chicks were sexed, wing-banded, and vaccinated for Infectious Bronchitis and Marek's disease. Beaks were kept intact. At approximately 17 weeks of age, the laying hens were put in a laying house. Hens were housed in the top two levels of each row. Five hens of the same cross and age were assigned to a cage. All hens within a cage had the same sire. No dam information was available. Limited interaction with back and side neighbours was possible.

The trait of interest, survival time, was defined as the number of days from the moment the hens were placed in the laying house until either death or the end of the experiment. Cages were checked daily to remove and record dead laying hens. Depending on the batch, the experiment ended 347 up to 435 days after placing the hens in the laying house. Conservatively, the maximum survival time of all batches was set at the survival time of the batch with the shortest experimental time span. We analysed two scenarios. In the first scenario, the maximum survival time of all eight batches was set at 347 days. From the eight available batches, six had an experimental time span equal or larger than 414 days. Potentially it would be more informative to delete the two batches with short experimental time span and to extend the maximum survival time of the six remaining batches. The phenotypic variance will increase and the percentage of censored records will decrease. Therefore, in the second scenario, the maximum survival time of those six batches was set at 414 days. Survival time with a maximum of 347 days (ST_{347}) and survival time with a maximum of 414 days (ST_{414}) are analysed in this study. Table 3.2 shows the number of phenotypes and (genotyped) sires per sire line when the dataset contained either eight batches censored at 347 days or six batches censored at 414 days.

Table 3.2 Number of phenotypes and (genotyped) sires for each dataset and sire line

	ST_{347}			ST_{414}		
	W1	W5	W6	W1	W5	W6
# Phenotypes	20 520	13 495	16 575	14 975	8 215	12 390
# Sires†	507	364	464	388	210	341
# Genotyped sires	342	352	450	252	202	329

† Both non-genotyped and genotyped sires

3.3 Methods

3.3.1 Variance components

Variance components were estimated for survival time at two different censoring-moments (347 and 414 days). Survival time in non-beak-trimmed group-housed birds is known to be influenced by the social interactions among birds (Craig and Muir, 1996; Muir, 1996; Ellen *et al.*, 2008, 2013, 2014; Peeters *et al.*, 2012). With

social interactions, a phenotype consists of the direct genetic (A_D) and environmental (E_D) effect of the individual itself (i), and the indirect genetic (A_I) and environmental (E_I) effects of its group mates (j):

$$P_i = A_{D_i} + \sum_{i \neq j}^{n-1} A_{I_j} + E_{D_i} + \sum_{i \neq j}^{n-1} E_{I_j} \quad [3.1]$$

where n is the number of individuals per cage (Griffing, 1967). In this particular case, all hens within a cage had the same sire. Lack of dam information required the use of a sire model. The genetic component in Equation 3.1 can be written in terms of the sire, dam and Mendelian sampling term:

$$P_i = \frac{1}{2}\left[A_{Sire_{D_i}} + \sum_{i \neq j}^{n-1} A_{Sire_{I_j}}\right] + \frac{1}{2}\left[A_{Dam_{D_i}} + \sum_{i \neq j}^{n-1} A_{Dam_{I_j}}\right]$$
$$+ MS_{D_i} + \sum_{i \neq j}^{n-1} MS_{I_j} + E_{D_i} + \sum_{i \neq j}^{n-1} E_{I_j}. \quad [3.2]$$

Because all animals within a cage have the same sire, $\sum_{i \neq j}^{n-1} A_{Sire_{I_j}}$ in Equation 3.2 can be replaced by $(n-1)A_{Sire_{I_j}}$:

$$P_i = \frac{1}{2}\left[A_{Sire_{D_i}} + (n-1)A_{Sire_{I_j}}\right] + \frac{1}{2}\left[A_{Dam_{D_i}} + \sum_{i \neq j}^{n-1} A_{Dam_{I_j}}\right]$$
$$+ MS_{D_i} + \sum_{i \neq j}^{n-1} MS_{I_j} + E_{D_i} + \sum_{i \neq j}^{n-1} E_{I_j}. \quad [3.3]$$

Because the sire's direct and indirect genetic effect are expressed in the same phenotype, the direct and indirect genetic parameters are completely confounded. However, the sire component can be rewritten in terms of the total genetic effect (A_T). From an animal breeding perspective, A_T is of interest, because it determines total response to selection (Bijma *et al.*, 2007). An animal's A_T consists of a direct and indirect component,

$$A_T = A_D + (n-1)A_I. \quad [3.4]$$

Using Equation 3.3 and 3.4, the sire's direct and indirect genetic effect can be replaced by its total genetic effect:

$$P_i = \frac{1}{2}A_{Sire_{T_i}} + \frac{1}{2}\left[A_{Dam_{D_i}} + \sum_{i \neq j}^{n-1} A_{Dam_{I_j}}\right]$$
$$+ MS_{D_i} + \sum_{i \neq j}^{n-1} MS_{I_j} + E_{D_i} + \sum_{i \neq j}^{n-1} E_{I_j}. \quad [3.5]$$

Hence, even though this experimental set-up does not allow the estimation of direct and indirect genetic effects separately, it does allow the estimation of total genetic effects when using a sire model.

The following sire model was used to estimate variance components:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_{\mathrm{T_{Sire}}}\mathbf{a}_{\mathrm{T_{Sire}}} + \mathbf{Vcage} + \mathbf{e}, \quad [3.6]$$

where $\mathbf{y}$ is a vector of individual records, $\mathbf{X}$ is an incidence matrix linking the individual records to fixed effects; $\mathbf{b}$ is a vector of fixed effects (a fixed effect for the cross, an interaction term for each laying house by row by level combination (also corrects for the batch effect), a fixed effect for the content of the surrounding cages (all three full or 1-2-3 empty) and a covariate for the average survival time in the surrounding cage); $\mathbf{Z}_{\mathrm{T_{Sire}}}$ is an incidence matrix linking the individual records to the $\mathrm{A_T}$'s of the sire; $\mathbf{a}_{\mathrm{T}}$ is a vector of $\mathrm{A_T}$'s; $\mathbf{V}$ is an incidence matrix linking the individual records to random cage effects; $\mathbf{cage}$ is a vector of random cage effects; and $\mathbf{e}$ is a vector of residuals.

Thus, the model has three random effects:

-The genetic sire effect; accounting for one quarter of the total genetic variance, where the total genetic variance consists of direct and indirect genetic (co)variances,

$$\sigma^2_{\mathrm{A_{T\_Sire}}} = \frac{1}{4}\sigma^2_{\mathrm{A_T}} = \frac{1}{4}\left(\sigma^2_{\mathrm{A_D}} + 2(\mathrm{n}-1)\sigma_{\mathrm{A_{DI}}} + (\mathrm{n}-1)^2\sigma^2_{\mathrm{A_I}}\right). \quad [3.7]$$

-The random cage effect; accounting for the non-genetic covariance among cage mates (Bergsma *et al.*, 2008) as well as the genetic covariance among cage mates that was not captured by the sire,

$$\sigma^2_{\mathrm{Cage}} = 2\sigma_{\mathrm{E_{DI}}} + (\mathrm{n}-2)\sigma^2_{\mathrm{E_I}} + \frac{3}{4}\left(2\sigma_{\mathrm{A_{DI}}} + (\mathrm{n}-2)\sigma^2_{\mathrm{A_I}}\right). \quad [3.8]$$

-The residual; accounting for the remaining non-genetic variance (Bijma, 2011b) as well as the remaining genetic variance that was not captured by the sire nor the cage variance,

$$\sigma^2_{\mathrm{E}} = \sigma^2_{\mathrm{E_D}} - 2\sigma_{\mathrm{E_{DI}}} + \sigma^2_{\mathrm{E_I}} + \frac{3}{4}\left(\sigma^2_{\mathrm{A_D}} - 2\sigma_{\mathrm{A_{DI}}} + \sigma^2_{\mathrm{A_I}}\right). \quad [3.9]$$

Analogous to the ordinary heritability, where the direct genetic variance is expressed relative to phenotypic variance, for social interaction traits, total heritable variance can be expressed relative to the phenotypic variance (Bergsma *et al.*, 2008):

$$\mathrm{T}^2 = \frac{\sigma^2_{\mathrm{A_T}}}{\sigma^2_{\mathrm{P}}}. \quad [3.10]$$

For each sire line, variance components as well as T^2's were compared for survival time censored at either 347 or 414 days.

3.3.2 H-matrix

To estimate variance components and breeding values, the blupf90-family programs were used (renumf90, airemlf90, preGSf90) (Misztal *et al.*, 2002). These programs use the single-step method as presented by Aguilar *et al.* (2010). The single-step method uses a relationship matrix (H-matrix) that combines the pedigree-based relationship matrix (A-matrix) and genomic relationship matrix (G-matrix) (Misztal *et al.*, 2009; Legarra *et al.*, 2009; Christensen and Lund, 2009). The inverse of the H-matrix has a simple form:

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}, \quad [3.11]$$

where $\mathbf{H}$ is the combined pedigree and genotype relationship matrix, $\mathbf{A}$ is the pedigree-based relationship matrix ($\mathbf{A}_{22}$ is the pedigree relationship matrix of genotyped animals only) and $\mathbf{G}$ is the genomic relationship matrix. To ensure that all matrices are invertible and to control bias, the genomic relationship matrix was adjusted to become compatible with the pedigree-based relationship matrix. $\mathbf{G}^{-1}$ in Equation 3.11 was replaced by $(\alpha\mathbf{G} + \beta\mathbf{A}_{22})^{-1}$, where α was set to 0.95 and β was set to 0.05, which are the default values in the preGSf90 software.

Prior to calculating the G-matrix, genotypes were subjected to quality control. Monomorphic SNPs, SNPs with a minor allele frequency (MAF) below 5%, a call rate below 90% and a departure from Hardy-Weinberg Equilibrium (HWE; maximum difference between observed and expected frequency of heterozygotes is 0.15) were excluded from further analyses. The departure from HWE was tested for autosomes and sex-chromosomes, since all genotypes were collected on sires only (all ZZ). On average, 37% of the 56 492 SNPs was retained after quality control (23 016 for W1, 17 621 for W5 and 23 463 for W6). The majority of SNPs was excluded because they were monomorphic (Table 3.3). This is because the SNP-chip was developed to be used in a wide range of broiler and layer lines. Even though more than half of the SNPs on the 60k SNP-chip were fixed within a line, these monomorphic SNPs were still informative for some of the other lines.

Table 3.3 Number of SNPs that failed quality control for each dataset and sire line

	ST_{347}			ST_{414}		
	W1	W5	W6	W1	W5	W6
Call rate	951	1 143	1 057	1 010	1 066	1 083
MAF†	33 472	38 485	32 685	33 274	38 645	32 794
Monomorphic	29 532	34 523	28 706	29 536	34 782	28 797
HWE††	88	13	5	108	71	45

† Minor allele frequency; include monomorphic SNPs

†† Hardy Weinberg Equilibrium

3.3.3 Accuracy of estimated breeding values

Through cross validation, the accuracy of estimated breeding values was calculated when:

(i) using genotypic information in addition to pedigree information (EBV *vs* GBV) and/or
(ii) extending the moment of censoring (ST_{347} *vs* ST_{414}).

The cross validation was performed by masking the daughter-phenotypes of 20 to 30 genotyped sires. To allow a fair comparison between ST_{347} and ST_{414}, the same sires were selected in both datasets. This was repeated 10 times, with no overlap in selected sires between repeats. Normally, to obtain the accuracy, the predicted breeding value of the sires with masked daughter-phenotype would be correlated to the mean daughters' performance, corrected for fixed effects, and divided by the accuracy of progeny testing. However, in the current dataset, 67 to 84% of the records are censored. This requires a slightly different approach when calculating the accuracy. Although no distinction can be made between censored animals, it is known that they outperformed the non-censored animals. To properly utilize this information, Ellen *et al.* (2010) proposed to rank corrected phenotypes, where animals with a censored phenotype would get the average phenotypic rank of all censored animals. For example, in an experiment with N animals, from which $N - n$ are censored, the non-censored animals receive rank 1 to n, while all censored animals receive the average rank of animal $n + 1$ up until animal N, *i.e.* a value of $[(n + 1 + N)]/2$ (Figure 3.1).

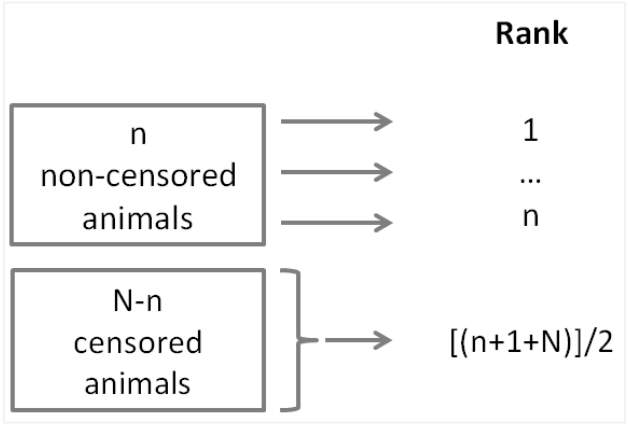

Figure 3.1 Ranking animals according to their corrected phenotypes while taking into account that a proportion of the animals has a censored phenotype

The predicted breeding value of the sires with masked daughter-phenotype was correlated to the mean daughters' rank and divided by the accuracy of progeny testing. The accuracy of progeny testing was calculated as:

$$\mathrm{Acc}_{\mathrm{Progeny\_testing}} = \sqrt{\frac{\frac{1}{4}\sigma^2_{\mathrm{A_T}}}{\mathrm{Var}(\overline{\mathrm{P}}_{\mathrm{Progeny}})}} \qquad [3.12]$$

where $\mathrm{Var}(\overline{\mathrm{P}}_{\mathrm{Progeny}})$ is the phenotypic variance of the mean progeny performance.

3.4 Results and discussion

3.4.1 Descriptive statistics

W1 had a higher mortality than W5 and W6, particularly in the first stage of the laying period (Figure 3.2). However, lines and batches were strongly confounded. Therefore, the underperformance of W1 might well be a batch-effect.

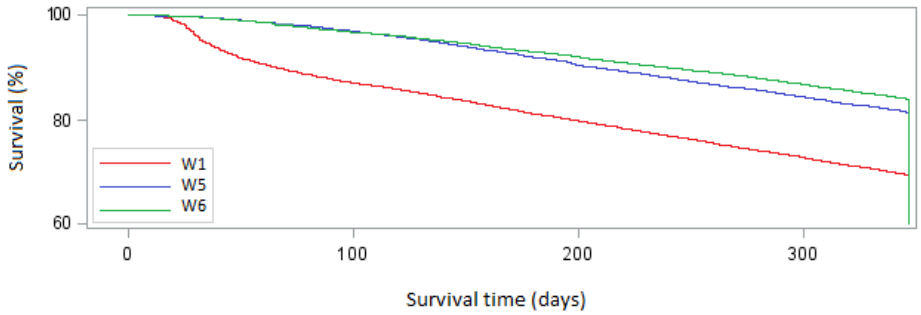

Figure 3.2 The Kaplan–Meier survival curve, showing survival (%) across time (days, with a maximum of 347) for the offspring of three sire lines

Delaying the moment of censoring from 347 to 414 days caused an average increase in mean survival time of 52 days (Table 3.4). Consequently, the phenotypic variance increased. In addition, the percentage of censored records decreased (Table 3.4).

Table 3.4 Mean survival time (days) and the percentage of censored records for two datasets and three sire lines

	ST_{347}			ST_{414}		
	W1	W5	W6	W1	W5	W6
Mean survival time	286	319	323	338	371	375
% Uncensored	31	19	16	33	23	21

3.4.2 Variance components

For ST_{347}, the total genetic variance relative to the phenotypic variance (T^2) was 0.30 ± 0.03 for crossbreds with a W1 sire, 0.16 ± 0.02 for crossbreds with a W5 sire and 0.12 ± 0.02 for crossbreds with a W6 sire (Table 3.5). The estimated T^2's are similar to the T^2's previously found in purebreds (0.16-0.19 (Ellen *et al.*, 2008)) and crossbreds (0.17-0.26 (Peeters *et al.*, 2012; Alemu *et al.*, submitted)).

By delaying the moment of censoring from 347 days to 414 days, the phenotypic variance increased by 46% for W1, 78% for W5 and 84% for W6. Because the genetic variance increased with approximately the same percentage, T^2 remained fairly stable (Table 3.5). These results are consistent with literature. Ellen *et al.* (2008) analysed survival time censored at 447 days in three White Leghorn purebred lines (W1, WB and WF). Peeters *et al.* (2012) analysed survival time censored at 398 days in a White Leghorn reciprocal cross (W1xWB and WBxW1). In order to make a proper comparison between purebred and crossbred performance, Peeters *et al.* (2012) re-analysed the W1 and WB purebred data, but this time censored at 398 days. When comparing $\sigma^2_{A_T}$, σ^2_P and T^2 for W1 and WB censored at either 398 or 447 days (Ellen *et al.*, 2008; Peeters *et al.*, 2012), the same observation was made. The phenotypic variance increased by 31% for W1 and 26% for WB. But, because the genetic variance increased with approximately the same percentage, T^2 stayed stable at 0.19 for W1 and only slightly increased from 0.15 to 0.16 for WB.

3.4.3 Accuracy of estimated breeding values

The accuracy of EBVs varied between 0.16 and 0.23 (Table 3.6). By using genotypic information in addition to pedigree information, all lines showed an increased accuracy. The accuracy of GBVs varied between 0.19 and 0.33 (Table 3.6). The increase in accuracy varied between 20 and 110%. This shows that the breeding

value estimation of genotyped sires without progeny information would benefit from genomics. This would help improve survival time in non-beak-trimmed group-housed laying hens. Alemu *et al.*, (submitted) calculated the accuracy of EBVs and GBVs for survival time in brown layers, with a similar experimental set-up. The accuracy of EBVs varied between 0.25 and 0.35, while the accuracy of GBVs varied between 0.34 and 0.48 (Alemu *et al.*, submitted). By using genotypic information in addition to pedigree information, accuracy increased between 35 and 40%. Overall, the accuracy in brown layers was higher than in white layers. This could be attributed to the lower survival in brown layers, which resulted in a smaller proportion of censored animals and a more informative dataset.

By extending the moment of censoring from 347 to 414 days, W1 showed a decreased accuracy, both for EBVs and GBVs. The phenotypic variance increased by 46%. However, this did not outweigh the loss of 27% of data. In contrast, W5 and W6 showed a stable accuracy for EBVs and an increased accuracy for GBVs. Both lines were more heavily censored than W1. The phenotypic variance increased by ~81% in these two lines. This benefit slightly outweighed the loss of ~32% of data in these two lines. However, differences were small and standard errors large. Despite that the results are inconclusive, it is clear that there is an intermediate optimum, as the two extremes (small dataset and high phenotypic variation *vs* large dataset and low phenotypic variation) will not result in accurate breeding values.

3.5 Conclusions

For survival time in non-beak-trimmed group-housed birds, T^2 varied from 0.12 to 0.30. Genetic improvement can, therefore, help to decrease mortality due to feather pecking and cannibalism. By using genotypic information in addition to pedigree information, the accuracy of the estimated total breeding values increased by 20 up to 110% for sires without progeny information, showing the added value of genomic selection. Moreover, it could be beneficial to extend the moment of censoring to create more phenotypic variation. However, if this implies deleting large batches with a shorter experimental time span, the benefit could be lost due to the loss of power.

Table 3.5 Variance components and T^2 for survival time for two datasets and three sire lines

	ST_{347}			ST_{414}		
	W1	W5	W6	W1	W5	W6
$\sigma^2_{A_{T_{Sire}}}$	777 ± 82	195 ± 30	132 ± 21	1 102 ± 132	301 ± 60	259 ± 45
σ^2_{Cage}	1 189 ± 71	164 ± 34	256 ± 30	1 561 ± 117	308 ± 77	436 ± 62
σ^2_E	8 368 ± 92	4 471 ± 61	3 985 ± 49	12 378 ± 160	7 981 ± 139	7 370 ± 105
σ^2_P	10 334 ± 123	4 830 ± 62	4 373 ± 50	15 041 ± 207	8 590 ± 140	8 065 ± 107
T^2	0.30 ± 0.03	0.16 ± 0.02	0.12 ± 0.02	0.29 ± 0.03	0.14 ± 0.03	0.13 ± 0.02

Table 3.6 Accuracy of EBVs (obtained with an A-matrix) and GBVs (obtained with an H-matrix) for two datasets and three sire line

	ST_{347}			ST_{414}		
	W1	W5	W6	W1	W5	W6
Accuracy EBV	0.23 ± 0.06	0.16 ± 0.07	0.19 ± 0.05	0.16 ± 0.06	0.16 ± 0.07	0.19 ± 0.05
Accuracy GBV	0.28 ± 0.06	0.23 ± 0.07	0.29 ± 0.05	0.19 ± 0.06	0.33 ± 0.07	0.30 ± 0.05

3.6 Acknowledgements

This research was financially supported by the Netherlands Organization for Scientific Research (NWO) and coordinated by the Dutch Technology Foundation (STW). Institut de Sélection Animale B.V., a Hendrix Genetics company, provided the data and was closely involved in this research. The authors thank Daniela Lino and Sophie Eaglen for their valuable input.

4

Using pooled data to estimate variance components and breeding values for traits affected by social interactions

K. Peeters[1], E. D. Ellen[1], P. Bijma[1]

[1] Animal Breeding and Genomics Centre, Wageningen University, 6700 AH Wageningen, The Netherlands

GSE (2013) 45: 27

Abstract

Through social interactions, individuals affect one another's phenotype. In such cases, an individual's phenotype is affected by the direct (genetic) effect of the individual itself and the indirect (genetic) effects of the group mates. Using data on individual phenotypes, direct and indirect genetic (co)variances can be estimated. Together, they compose the total genetic variance that determines a population's potential to respond to selection. However, it can be difficult or expensive to obtain individual phenotypes. Phenotypes on traits such as egg production and feed intake are, therefore, often collected on group level. In this study, we investigated whether direct, indirect and total genetic variances, and breeding values can be estimated from pooled data (pooled by group). In addition, we determined the optimal group composition, *i.e.* the optimal number of families represented in a group to minimise the standard error of the estimates. First, all research questions were answered by theoretical derivations. Second, a simulation study was conducted to investigate the estimation of variance components and optimal group composition. Third, individual and pooled survival time records on 12 944 purebred laying hens were analysed to investigate the estimation of breeding values and response to selection. Through theoretical derivations and simulations, we showed that the total genetic variance can be estimated from pooled data, but the underlying direct and indirect genetic (co)variances cannot. Moreover, we showed that the most accurate estimates are obtained when group members belong to the same family. Data analyses on survival time records showed that the correlation between the estimated total breeding values obtained from individual and pooled data was surprisingly close to one. This indicates that, for survival time in purebred laying hens, loss in response to selection will be small when using pooled instead of individual data. In conclusion, using pooled data, the total genetic variance and breeding values can be estimated, but the underlying genetic components cannot. The most accurate estimates are obtained when group members belong to the same family.

Key words: Social interactions, indirect genetic effects, pooled data, survival time, laying hens

4.1 Background

Group housing is common practice in most livestock farming systems. Previous studies have shown that group-housed animals can substantially affect one another's phenotype through social interactions (Craig, 1982; Muir, 2005; Van Vleck *et al.*, 2007; Chen *et al.*, 2008; Ellen *et al.*, 2008; Chen *et al.*, 2009; Duijvesteijn *et al.*, 2012; Peeters *et al.*, 2012; Muir *et al.*, 2013). The heritable effect of an individual on its own phenotype is known as the direct genetic effect, while the heritable effect of an individual on the phenotype of a group mate is known as the social, associative or indirect genetic effect (Willham, 1963; Griffing, 1967; Moore *et al.*, 1997; Wolf *et al.*, 1998; Bijma *et al.*, 2007). Both direct and indirect genetic effects determine a population's potential to respond to selection, *i.e.* the total genetic variance (Willham, 1963; Griffing, 1967; Moore *et al.*, 1997; Wolf *et al.*, 1998; Muir, 2005; Bijma *et al.*, 2007). Selection experiments in laying hens and quail (Craig,1982; Muir, 2005; Muir *et al.*, 2013], and variance component estimates in laying hens, quail, beef cattle and pigs (Van Vleck *et al.*, 2007; Chen *et al.*, 2008; Ellen *et al.*, 2008; Chen *et al.*, 2009; Duijvesteijn *et al.*, 2012; Peeters *et al.*, 2012; Muir *et al.*, 2013) have shown that indirect genetic effects can contribute substantially to the total genetic variation in agricultural populations.

Direct, indirect and total genetic variances can be estimated from individual data. However, it can be difficult or expensive to obtain individual phenotypes on certain traits, *e.g.* egg production and feed intake. Alternatively, data can be obtained on group level, resulting in pooled records. However, pooling data reduces the number of data points. Moreover, multiple animals influence each data point, increasing the complexity of the data. Although there is an obvious loss of power, previous studies have shown that pooled data can be used to estimate direct genetic variances for traits not affected by social interactions (Olson *et al.*, 2006; Biscarini *et al.*, 2008; Biscarini *et al.*, 2010b). However, with social interactions, indirect genetic effects emerge and the complexity of the data increases further. It is unclear whether pooled data are still informative in these situations. Therefore, the main objective of this study was to determine whether pooled data can be used to estimate direct, indirect and total genetic variances, and breeding values for traits affected by social interactions. In addition, optimal group composition was determined, *i.e.* the optimal number of families represented in a group to minimise the standard error of the estimates.

4.2 Methods

This study was performed in three steps. First, all research questions were answered by theoretical derivations. Second, a simulation study was conducted to investigate the estimation of variance components and optimal group composition. Third, individual and pooled survival time records on 12 944 purebred laying hens were analysed to investigate the estimation of breeding values and response to selection.

Table 4.1 lists the main symbols and their meaning.

4.2.1 Theory

Variance components and breeding value estimation

In this section, we examined whether direct, indirect and total genetic variances, and breeding values can be estimated from pooled data.

With social interactions, an individual phenotype consists of the direct genetic (A_D) and environmental (E_D) effects of the individual itself (i), and the indirect genetic (A_I) and environmental (E_I) effects of its group mates (j):

$$P_i = A_{D_i} + E_{D_i} + \sum_{i \neq j}^{n-1} A_{I_j} + \sum_{i \neq j}^{n-1} E_{i_j}, \quad [4.1]$$

where n is the number of individuals per group (Griffing, 1967). From an animal breeding perspective, the total breeding value (A_T) is of interest, because it determines total response to selection. An animal's A_T consists of a direct and indirect component:

$$A_{T_i} = A_{D_i} + (n-1)A_{I_i}. \quad [4.2]$$

where A_D is expressed in the phenotype of the animal itself and A_I is expressed in the phenotype of each group mate.

A pooled record (P^*) consists of the individual phenotypes of all group members (k):

$$P^* = \sum_{k=1}^{n} P_k. \quad [4.3]$$

It follows from Equations 4.1 and 4.3 that, with social interactions, a pooled record consists of the A_D and E_D of each group member, as well as their A_I and E_I that are expressed $n-1$ times:

$$P^* = \sum_{k=1}^{n} \left[A_{D_k} + E_{D_k} + (n-1)\left(A_{I_k} + E_{I_k}\right)\right] \quad [4.4]$$

Because an animal's A_D and A_I are expressed in the same pooled record, the direct **Z**-matrix that links pooled phenotypes to A_D's and the indirect **Z**-matrix that links pooled phenotypes to A_I's are completely confounded (as shown in *Appendix A*).

Table 4.1 Notation key

Symbol	Meaning
$i - j$	Focal individual - Group mates of the focal individual
A_D	Direct genetic effect \ Direct breeding value
A_I	Indirect genetic effect \ Indirect breeding value
A_T	Total genetic effect \ Total breeding value
E_D	Direct environmental effect
E_I	Indirect environmental effect
$\sigma^2_{A_D}$	Direct genetic variance
$\sigma_{A_{DI}}$	Direct-indirect genetic covariance
$\sigma^2_{A_I}$	Indirect genetic variance
$\sigma^2_{A_T}$	Total genetic variance
σ^2_{Cage}	Cage variance
σ^2_E	Error variance
σ^2_P	Phenotypic variance
$\sigma^2_{E^*}$	Pooled error variance
$\sigma^2_{P^*}$	Pooled phenotypic variance
h^2	Direct genetic variance relative to phenotypic variance \ Heritability
T^2	Total genetic variance relative to phenotypic variance
σ^2_z	Full variance
σ^2_b	Between-family variance
σ^2_w	Within-family variance
r	Relatedness within a family
N	Number of families
m	Number of records per family
o	Family size
n	Group size
^	Hat, denotes estimated values

Consequently, direct and indirect (co)variances, and breeding values cannot be estimated from pooled data.

It follows from Equations 4.2 and 4.4 that, with social interactions, a pooled record contains the total genetic effect of each group member:

$$P^* = \sum_{k=1}^{n} \left[A_{T_k} + E_k\right]. \qquad [4.5]$$

Equation 4.5 shows strong similarities with:

$$P^* = \sum_{k=1}^{n} \left[A_{D_k} + E_k\right], \qquad [4.6]$$

which shows the content of a pooled record when social interactions do not occur. Previous studies have shown that pooled data can be used to estimate direct genetic variances ($\sigma^2_{A_D}$) and direct breeding values for traits that are not affected by social interactions (Olson *et al.*, 2006; Biscarini *et al.*, 2008; Biscarini *et al.*, 2010b). Similarly, pooled data can be used to estimate total genetic variances ($\sigma^2_{A_T}$) and total breeding values for traits that are affected by social interactions.

Optimal group composition

In this section, the standard error (s.e.) of $\widehat{\sigma}^2_{A_T}$ is derived for three experimental designs that differ with respect to group composition, *i.e.* group members belonged to either one, two or n families. The s.e. of an estimate of the genetic variance depends on the between- (σ^2_b) and within-family variance (σ^2_w), the relatedness within a family (r), the number of families (N), and the number of records per family (m) (Lynch and Walsh, 1998):

$$\text{s.e.}\left(\widehat{\sigma}^2_A\right) \approx \frac{1}{r}\sqrt{\frac{2}{N-1}\left[\sigma^4_b + \frac{2\sigma^2_b\sigma^2_w}{m} + \frac{\sigma^4_w}{m(m-1)}\right]}. \qquad [4.7]$$

Analysis of variance was used to derive σ^2_b and σ^2_w for each design (see *Appendix B* for derivation).

The s.e. of $\widehat{\sigma}^2_{A_T}$ differs between experimental designs, because the group composition changes the within-family variance and the number of records per family (Table 4.2). On the one hand, the within-family variance decreases when the number of families per group decreases, causing a strong decrease in s.e.. On the other hand, the number of records per family decreases when the number of families per group decreases, causing a slight increase in s.e.. Overall, to obtain the most accurate estimate of $\sigma^2_{A_T}$, group members should belong to the same family.

The only exception is when family size (o) equals group size (n). In this case, there is only one record per family and $\sigma^2_{A_T}$ would not be estimable.
Ideally, group members should be full-sibs rather than half-sibs, since an increase in relatedness causes a decrease in the s.e. of $\widehat{\sigma}^2_{A_T}$.

Table 4.2 Within-family variance (σ^2_w) and number of records per family (m) for three group compositions

	σ^2_w	m
One family	$\frac{1}{n}\left[\sigma^2_{P_D} + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma^2_{P_I} + (n-1)r\sigma^2_{A_T}\right] - r\sigma^2_{A_T}$	$\frac{o}{n}$
Two families	$\frac{4}{n}\left[\sigma^2_{P_D} + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma^2_{P_I} + \left(\frac{n}{2}-1\right)r\sigma^2_{A_T}\right] - r\sigma^2_{A_T}$	$\frac{2o}{n}$
n families	$n\left[\sigma^2_{P_D} + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma^2_{P_I}\right] - r\sigma^2_{A_T}$	o

r, N, n, o and σ^2_b do not differ between group compositions

4.2.2 Simulation

To validate the theoretical derivations, a simulation study was conducted in R v2.12.2 (Venables *et al.*, 2011). A base population of 500 sires and 500 dams was simulated. Each animal in the base population was assigned a direct and indirect breeding value, drawn from $N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_{A_D} & \sigma_{A_{DI}}\\ \sigma_{A_{DI}} & \sigma^2_{A_I}\end{bmatrix}\right)$. The $\sigma^2_{A_D}$ and $\sigma^2_{A_I}$ were set to 1.00, and $\sigma_{A_{DI}}$ was set to −0.50, 0.00 or 0.50. Each sire was randomly mated to a single dam, resulting in 12 offspring per mating and a total of 6 000 simulated offspring. For each offspring, direct and indirect breeding values were obtained as $A_D = \frac{1}{2}A_{D_{Sire}} + \frac{1}{2}A_{D_{Dam}} + MS_D$ and $A_I = \frac{1}{2}A_{I_{Sire}} + \frac{1}{2}A_{I_{Dam}} + MS_I$, where the direct and indirect Mendelian sampling terms were drawn from $N\left(\begin{bmatrix}0\\0\end{bmatrix}, \frac{1}{2}\begin{bmatrix}\sigma^2_{A_D} & \sigma_{A_{DI}}\\ \sigma_{A_{DI}} & \sigma^2_{A_I}\end{bmatrix}\right)$. Each offspring was also assigned a direct and an indirect environmental value, drawn from $N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_{E_D} & \sigma_{E_{DI}}\\ \sigma_{E_{DI}} & \sigma^2_{E_I}\end{bmatrix}\right)$. The $\sigma^2_{E_D}$ and $\sigma^2_{E_I}$ were set to 2.00, and $\sigma_{E_{DI}}$ was set to −1.00, 0.00 or 1.00. Animals were placed in groups of four. Depending on the scenario, group members belonged to one, two or four families.

Individual phenotypes were obtained by summing the direct and indirect genetic and environmental components according to Equation 4.1. Pooled records were obtained by summing individual phenotypes according to Equation 4.3. Seven scenarios were simulated, which differed in $\sigma_{A_{DI}}$, $\sigma_{E_{DI}}$ or group composition (Table 4.3). For each scenario, 100 replicates were produced.

Table 4.3 Scenarios used to simulate data

	Scenario[§]	$\sigma_{A_{DI}}$	$\sigma_{E_{DI}}$	Group composition
Reference scenario	1	0.00	0.00	Four families
Different $\sigma_{A_{DI}}$	2	−0.50	0.00	Four families
	3	0.50	0.00	Four families
Different $\sigma_{E_{DI}}$	4	0.00	−1.00	Four families
	5	0.00	1.00	Four families
Different group compositions	6	0.00	0.00	Two families
	7	0.00	0.00	One family

§ $\sigma^2_{A_D}$ and $\sigma^2_{A_I}$ were set to 1.00; $\sigma^2_{E_D}$ and $\sigma^2_{E_I}$ were set to 2.00

Based on the previous section, expectations are that the use of a direct-indirect animal model for pooled data will fail to differentiate between direct and indirect genetic effects, while the use of a traditional animal model for pooled data will yield estimates of $\sigma^2_{A_T}$. To validate these theoretical predictions, both models were run.

First, the simulated pooled records were analysed with the following direct-indirect animal model in ASReml v3.0 (Gilmour *et al.*, 2009):

$$\mathbf{y}^* = \boldsymbol{\mu}^* + \mathbf{Z}^*_D\mathbf{a}_D + \mathbf{Z}^*_I\mathbf{a}_I + \mathbf{e}^*, \quad [4.8]$$

where $\mathbf{y}^*$ is a vector that contains pooled records (P^*); $\boldsymbol{\mu}^*$ is a vector that contains the pooled mean; $\mathbf{Z}^*_D$ is an incidence matrix linking the pooled records to A_D's (each pooled record was linked to the A_D's of the four group members); $\mathbf{a}_D$ is a vector that contains A_D's; $\mathbf{Z}^*_I$ is an incidence matrix linking the pooled records to A_I's (each pooled record was linked to the A_I's of the four group members); $\mathbf{a}_I$ is a vector that contains A_I's; and $\mathbf{e}^*$ is a vector that contains residuals.

Second, the simulated pooled records were analysed with the following traditional animal model in ASReml v3.0:

$$\mathbf{y}^* = \boldsymbol{\mu}^* + \mathbf{Z}^*\mathbf{a} + \mathbf{e}^* \quad [4.9]$$

where $\mathbf{y}^*$, $\boldsymbol{\mu}^*$ and $\mathbf{e}^*$ are as explained above; $\mathbf{Z}^*$ is an incidence matrix linking the pooled records to A's (each pooled record was linked to the A's of the four group members); and $\mathbf{a}$ is a vector that contains A's.

Based on the previous section, expectations are that the most accurate prediction of $\sigma^2_{A_T}$ will be obtained when group members belong to the same family. To validate this theoretical prediction, the predicted s.e. of $\widehat{\sigma}^2_{A_T}$ was compared to (i) the standard deviation (s.d.) of 100 estimates of $\sigma^2_{A_T}$ ($\widehat{\sigma}^2_{A_T}$'s reported by ASReml) and (ii) the mean of 100 s.e.'s of $\widehat{\sigma}^2_{A_T}$ (s.e.'s reported by ASReml) for three group compositions (scenarios 1, 6 and 7 of Table 4.3).

4.2.3 Data analyses

The dataset was part of the pre-existing database of Hendrix Genetics (The Netherlands) and contained routinely collected data for breeding value estimation. Animal Care and Use Committee approval was therefore not required.

To validate the theoretical derivations and to gain insight into response to selection, individual and pooled data on survival time in purebred laying hens (*Gallus gallus*) were analysed. Survival time in group-housed laying hens is a well-known example of a trait affected by social interactions, since a bird's chance to survive depends on the feather pecking and cannibalistic behaviour of its group mates. Ellen *et al.* (2008) used individual survival time data on three purebred lines to estimate direct and indirect genetic (co)variances. Large and statistically significant indirect genetic effects were found in two out of three purebred lines. In the current study, we used data from those two lines. Data were provided by the Institut de Sélection Animale B.V., the layer breeding division of Hendrix Genetics. Data on 13 192 White Leghorn layers were provided of which 6 276 were of line W1 and 6 916 were of line WB.

At the age of 17 weeks, the hens were placed in two laying houses. The laying houses consisted of four or five double rows, and each row consisted of three levels. Interaction with neighbours on the back of the cage was possible, but interaction with neighbours on the side was prevented. Four hens of the same purebred line were randomly assigned to each cage. Hens were not beak-trimmed. Further details on housing conditions and management are in Ellen *et al.* (2008).

The individual phenotype was defined as the number of days from the start of the laying period until either death or the end of the experiment, with a maximum of 398 days. The individual phenotypes were summed per cage to obtain pooled records. If one individual phenotype was missing, the entire cage was omitted from the analysis. The final dataset contained records on 6 092 W1 and 6 852 WB hens.

To obtain the direct, indirect and total genetic parameters for survival time, the individual phenotypes were analysed with the following direct-indirect animal model in ASReml v3.0:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_{\mathrm{D}}\mathbf{a}_{\mathrm{D}} + \mathbf{Z}_{\mathrm{I}}\mathbf{a}_{\mathrm{I}} + \mathbf{Vcage} + \mathbf{e} \quad [4.10]$$

where $\mathbf{y}$ is a vector that contains individual phenotypes; $\mathbf{X}$ is an incidence matrix linking the individual phenotypes to fixed effects; $\mathbf{b}$ is a vector that contains fixed effects, which included an interaction term for each laying house by row by level combination, an effect for the content of the back cage (full/empty) and a covariate for the average number of survival days in the back cage; $\mathbf{Z}_{\mathrm{D}}$ is an incidence matrix linking the individual phenotypes to A_{D}'s; $\mathbf{a}_{\mathrm{D}}$ is a vector that contains A_{D}'s; $\mathbf{Z}_{\mathrm{I}}$ is an incidence matrix linking the individual phenotypes to A_{I}'s; $\mathbf{a}_{\mathrm{I}}$ is a vector that contains A_{I}'s; $\mathbf{V}$ is an incidence matrix linking the individual phenotypes to random cage effects; **cage** is a vector that contains random cage effects (to account for the non-genetic covariance among phenotypes of cage members (Bergsma *et al.*, 2008)); and $\mathbf{e}$ is a vector that contains residuals. This model yields estimates of $\sigma^2_{\mathrm{A_D}}$, $\sigma_{\mathrm{A_{DI}}}$ and $\sigma^2_{\mathrm{A_I}}$, from which $\widehat{\sigma}^2_{\mathrm{A_T}}$ can be calculated. Similarly, it yields estimates of A_{D}'s and A_{I}'s, from which $\widehat{\mathrm{A}}_{\mathrm{T}}$'s can be calculated. To improve a trait, animals should be selected based on their $\widehat{\mathrm{A}}_{\mathrm{T}}$, since $\sigma^2_{\mathrm{A_T}}$ determines a population's potential to respond to selection.

Alternatively, a traditional animal model can be used to analyse individual or pooled data. A traditional animal model on individual data only yields estimates of $\sigma^2_{\mathrm{A_D}}$ and A_{D}'s. A traditional model on pooled data is expected to yield estimates of $\sigma^2_{\mathrm{A_T}}$ and A_{T}'s, but not of $\sigma^2_{\mathrm{A_D}}$ and A_{D}'s. To validate this theoretical prediction, these traditional models were also run.

First, the individual phenotypes were analysed with the following traditional (direct) animal model in ASReml v3.0:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_{\mathrm{D}}\mathbf{a}_{\mathrm{D}} + \mathbf{Vcage} + \mathbf{e} \quad [4.11]$$

where $\mathbf{y}$, $\mathbf{X}$, $\mathbf{b}$, $\mathbf{Z}_{\mathrm{D}}$, $\mathbf{a}_{\mathrm{D}}$, $\mathbf{V}$, **cage** and $\mathbf{e}$ are as explained above.

Second, the pooled records were analysed with the following traditional animal model in ASReml v3.0:

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}^* + \mathbf{Z}^*\mathbf{a} + \mathbf{e}^* \quad [4.12]$$

where $\mathbf{y}^*$ is a vector that contains pooled records (P^*); $\mathbf{X}^*$ is an incidence matrix linking the pooled records to fixed effects; $\mathbf{b}^*$ is a vector that contains fixed effects (the same fixed effects as mentioned above); $\mathbf{Z}^*$ is an incidence matrix linking the pooled records to A's (each pooled record was linked to the A's of the four group members); $\mathbf{a}$ is a vector that contains A's; and $\mathbf{e}^*$ is a vector that contains residuals.

The estimated variance components and breeding values of all three models were compared. In addition, we calculated the loss in response to selection that would occur when applying a traditional model to individual or pooled data instead of a direct-indirect model to individual data. The direct-indirect model applied to individual data yielded estimates of $\sigma^2_{\mathrm{A_T}}$ and $\mathrm{A_T}$'s. Based on their $\widehat{\mathrm{A}}_\mathrm{T}$, 250 animals were selected and the corresponding response to selection was calculated. Similarly, for the two traditional animal models, 250 animals were selected based on their $\widehat{\mathrm{A}}_\mathrm{D}$ (obtained from individual data) and $\widehat{\mathrm{A}}$ (obtained from pooled data). Once the top 250 animals were selected, their $\widehat{\mathrm{A}}_\mathrm{T}$ (obtained from individual data) was used to calculate the total response to selection. Then, the loss in total response to selection was calculated.

4.3 Results and discussion

4.3.1 Simulation

The direct-indirect animal model on pooled records failed to converge, confirming that direct and indirect (co)variances cannot be estimated from pooled data. The traditional animal model on pooled records yielded estimates of σ^2_A and $\sigma^2_{\mathrm{E}^*}$. These estimates did not differ significantly from the true $\sigma^2_{\mathrm{A_T}}$ and $\sigma^2_{\mathrm{E}^*}$ (Table 4.4), where

$$\sigma^2_{\mathrm{A_T}} = \sigma^2_{\mathrm{A_D}} + 2(\mathrm{n}-1)\sigma_{\mathrm{A_{DI}}} + (\mathrm{n}-1)^2\sigma^2_{\mathrm{A_I}} \quad [4.13]$$

(derived by Bijma *et al.* (2007)) and

$$\sigma^2_{\mathrm{E}^*} = \mathrm{n}\left[\sigma^2_{\mathrm{E_D}} + 2(\mathrm{n}-1)\sigma_{\mathrm{E_{DI}}} + (\mathrm{n}-1)^2\sigma^2_{\mathrm{E_I}}\right] \quad [4.14]$$

(analogous to what was found by Biscarini *et al.* (2010b)).

Table 4.4 True and estimated $\sigma^2_{A_T}$ and $\sigma^2_{E^*}$ for five scenarios

	Scenario[§]	$\sigma^2_{A_T}$ [§§]	$\overline{\hat{\sigma}^2_A \pm \text{s.e.}}$	$\sigma^2_{E^*}$ [§§§]	$\overline{\hat{\sigma}^2_{E^*} \pm \text{s.e.}}$
$\sigma_{A_{DI}} = 0.00$ $\sigma_{E_{DI}} = 0.00$	1	10.00	10.10 ± 1.85	80.00	80.56 ± 6.69
$\sigma_{A_{DI}} = -0.50$ $\sigma_{E_{DI}} = 0.00$	2	7.00	7.43 ± 1.59	80.00	79.29 ± 6.08
$\sigma_{A_{DI}} = 0.50$ $\sigma_{E_{DI}} = 0.00$	3	13.00	13.05 ± 2.12	80.00	80.32 ± 7.30
$\sigma_{A_{DI}} = 0.00$ $\sigma_{E_{DI}} = -1.00$	4	10.00	9.70 ± 1.54	56.00	56.54 ± 5.24
$\sigma_{A_{DI}} = 0.00$ $\sigma_{E_{DI}} = 1.00$	5	10.00	9.81 ± 2.10	104.00	104.71 ± 8.03

§ $\sigma^2_{A_D}$ and $\sigma^2_{A_I}$ were set to 1.00; $\sigma^2_{E_D}$ and $\sigma^2_{E_I}$ were set to 2.00; group members belonged to four different families

§§$\sigma^2_{A_T} = \sigma^2_{A_D} + 2(n-1)\sigma_{A_{DI}} + (n-1)^2\sigma^2_{A_I}$

§§§ $\sigma^2_{E^*} = n\left[\sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$

Based on Equation 4.7, the s.e. of $\hat{\sigma}^2_{A_T}$ was predicted for three scenarios that differed in group composition, *i.e.* group members belonged to one, two or four families. The theoretical s.e. of $\hat{\sigma}^2_{A_T}$ was compared to (i) the s.d. of 100 estimates of $\sigma^2_{A_T}$ ($\hat{\sigma}^2_{A_T}$'s reported by ASReml) and (ii) the mean of 100 s.e.'s of $\hat{\sigma}^2_{A_T}$ (s.e.'s reported by ASReml) (Table 4.5). The theoretical s.e. of $\hat{\sigma}^2_{A_T}$ did not differ significantly from the values obtained by simulation. Moreover, as predicted, the most accurate estimate of $\sigma^2_{A_T}$ was obtained when group members belonged to the same family. In comparison, the s.e. of $\hat{\sigma}^2_{A_T}$ was twice as large when group members belonged to different families. This indicates that group composition is crucial when aiming to obtain accurate estimates.

Table 4.5 Theoretically predicted s.e($\widehat{\sigma}^2_{A_T}$),s.d.($\widehat{\sigma}^2_{A_T}$)[§] and $\overline{\text{s. e.}\left(\widehat{\sigma}^2_{A_T}\right)}$ [§§] for three group compositions

	Scenario[§§§]	s.e($\widehat{\sigma}^2_{A_T}$),	s.d.($\widehat{\sigma}^2_{A_T}$) ± s.d.	$\overline{\text{s. e.}\left(\widehat{\sigma}^2_{A_T}\right)}$ ± s.d.
Four families	1	1.88	2.01 ± 0.14	1.85 ± 0.13
Two families	6	1.30	1.23 ± 0.09	1.23 ± 0.08
One family	7	0.92	0.81 ± 0.06	0.92 ± 0.05

§ s.d.($\widehat{\sigma}^2_{A_T}$) based on 100 $\widehat{\sigma}^2_{A_T}$'s reported by ASReml

§§ $\overline{\text{s. e.}\left(\widehat{\sigma}^2_{A_T}\right)}$ based on 100 s.e.'s reported by ASReml

§§§ $\sigma^2_{A_D}$ and $\sigma^2_{A_I}$ were set to 1.00; $\sigma_{A_{DI}}$ was set to 0.00; $\sigma^2_{E_D}$ and $\sigma^2_{E_I}$ were set to 2.00; $\sigma_{E_{DI}}$ was set to 0.00

4.3.2 Data analyses

Table 4.6 shows the estimated variance components for individual survival time data analysed with a direct-indirect animal model and the estimated variance components for individual and pooled survival time data analysed with a traditional animal model. The direct-indirect animal model on individual data yielded estimates of $\sigma^2_{A_D}$, $\sigma_{A_{DI}}$ and $\sigma^2_{A_I}$. Based on these components, $\widehat{\sigma}^2_{A_T}$ was calculated (according to Equation 4.13). The traditional animal model on individual data yielded estimates of $\sigma^2_{A_D}$. The traditional animal model on pooled data yielded estimates of σ^2_A that closely resembled the estimates of $\sigma^2_{A_T}$ from individual data. The direct-indirect animal model on individual data also yielded estimates of σ^2_{Cage} and σ^2_E. As derived by Bergsma *et al.* (2008), $\widehat{\sigma}^2_{Cage}$ is an estimate of $2\sigma_{E_{DI}} + (n-2)\sigma^2_{E_I}$. As derived by Bijma (2011b), $\widehat{\sigma}^2_E$ is an estimate of $\sigma^2_{E_D} - 2\sigma_{E_{DI}} + \sigma^2_{E_I}$. As shown in Equation 4.14, $\widehat{\sigma}^2_{E^*}$ is an estimate of $n\left[\sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$. Consequently, the $\widehat{\sigma}^2_{Cage}$ and $\widehat{\sigma}^2_E$ from the direct-indirect animal model on individual data should sum to the $\widehat{\sigma}^2_{E^*}$ from the traditional animal model on pooled data. More precisely:

$$\widehat{\sigma}^2_{E^*} = n^2\widehat{\sigma}^2_{Cage} + n\widehat{\sigma}^2_E. \qquad [4.15]$$

The expected $\widehat{\sigma}^2_{E^*}$, calculated based on the $\widehat{\sigma}^2_{Cage}$ and $\widehat{\sigma}^2_E$ from the direct-indirect animal model on individual data, and the $\widehat{\sigma}^2_{E^*}$ from the traditional animal model on pooled data closely resembled each other.

Table 4.6 Estimated variance components with s.e. from individual and pooled data on survival time in laying hens

	W1	WB
Direct-indirect animal model on individual data		
$\sigma^2_{A_D}$	705 ± 171	1 404 ± 301
$\sigma_{A_{DI}}$	59 ± 61	−162 ± 105
$\sigma^2_{A_I}$	104 ± 41	232 ± 72
σ^2_{Cage}	799 ± 166	1 191 ± 238
σ^2_E	7 980 ± 210	12 675 ± 365
$\sigma^2_{A_T}$ §	1 996 ± 640	2 521 ± 842
Expected $\sigma^2_{E^*}$ §§	44 700 ± 2 526	69 752 ± 3 513
Traditional (direct) animal model on individual data		
$\sigma^2_{A_D}$	677 ± 165	1 522 ± 317
σ^2_{Cage}	1 096 ± 127	1 443 ± 186
σ^2_E	8 002 ± 205	13 008 ± 338
Traditional animal model on pooled data		
σ^2_A	1 979 ± 643	2 521 ± 845
$\sigma^2_{E^*}$	44 750 ± 2 538	69 750 ± 3 519

§ In groups of four, $\sigma^2_{A_T}$ equals $\sigma^2_{A_D} + 6\sigma_{A_{DI}} + 9\sigma^2_{A_I}$
§§ In groups of four, $\sigma^2_{E^*}$ equals $16\sigma^2_{Cage} + 4\sigma^2_E$

Table 4.6 does not show heritability estimates. Where the classical heritability (h^2) is used to express $\sigma^2_{A_D}$ relative to the phenotypic variance (σ^2_P), T^2 is used to express $\sigma^2_{A_T}$ relative to σ^2_P (Bergsma *et al.*, 2008). Comparing values of T^2 obtained from individual and pooled data would be misleading because they are not expected to be similar. $\sigma^2_{P^*}$ cannot simply be divided by the number of group members to obtain σ^2_P. When group members are unrelated,

$$\sigma^2_P = \sigma^2_{A_D} + (n-1)\sigma^2_{A_I} + \sigma^2_{E_D} + (n-1)\sigma^2_{E_I} \quad [4.16]$$

and

$$\begin{aligned}\sigma^2_{P^*} &= n\sigma^2_{A_T} + \sigma^2_{E^*} \\ &= n[\sigma^2_{A_D} + 2(n-1)\sigma_{A_{DI}} + (n-1)^2\sigma^2_{A_I} + \\ &\quad \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}]\end{aligned} \quad [4.17]$$

The non-proportional increase of the phenotypic variance does not enable a meaningful comparison between values of T^2 obtained from individual and pooled data.

In conclusion, when group members are unrelated, a traditional animal model on individual data yields estimates of $\sigma^2_{A_D}$, while a traditional animal model on pooled data yields estimates of $\sigma^2_{A_T}$. Moreover, the estimated cage and error variances from a direct-indirect animal model on individual data sum to the pooled error variance from a traditional animal model on pooled data (Equation 4.15). This result could explain the 'inconsistencies' found by Biscarini *et al.* (2008), who assumed that a traditional animal model on individual and pooled data should yield the same genetic variance. Moreover, Biscarini *et al.* (2008) expected to find a pooled error variance that is four times larger than the individual error variance. For body weight at the age of 19 and 27 weeks, these expectations were met. For body weight at the age of 43 and 51 weeks, however, the genetic variance estimated from pooled data was smaller than expected, while the pooled error variance was larger than expected. Biscarini *et al.* (2008) mentions the emergence of competition effects as a possible cause. We indeed expect to find indirect genetic effects when the individual data on body weight at the age of 43 and 51 weeks were reanalysed with a direct-indirect animal model. Using Equations 4.13 and 4.15, the estimated variance components from individual data would resemble the estimated variance components from pooled data.

The regression coefficients of $\widehat{A}_D$'s obtained from individual data on the $\widehat{A}$'s obtained from pooled data strongly deviated from one (0.36 ± 0.01 for W1; 0.39 ± 0.01 for WB). The regression coefficients of $\widehat{A}_T$'s obtained from individual data on the $\widehat{A}$'s obtained from pooled data were close to, and not significantly different from, one (1.00 ± 0.01 for W1; 1.00 ± 0.01 for WB). This indicates that the $\widehat{A}$'s obtained from pooled data are unbiased estimates of the $\widehat{A}_T$'s obtained from individual data. Table 4.7 shows Spearman correlation coefficients between $\widehat{A}_D$'s and $\widehat{A}_T$'s obtained from individual data, and the $\widehat{A}$'s obtained from pooled data. The Spearman correlation coefficients between the $\widehat{A}_T$'s obtained from individual data and the $\widehat{A}$'s obtained from pooled data were close to, but significantly different from, one. This indicates only a minor loss in the accuracy of $\widehat{A}_T$'s when using pooled instead of individual data, which will be reflected in a minor loss in response to selection when using pooled instead of individual data.

Table 4.7 Spearman correlation coefficients between $\widehat{A}_D$'s and $\widehat{A}_T$'s obtained from individual data, and A's, with s.e., obtained from pooled data on survival time in laying hens

	A_D	A_T	A
A_D		0.513 ± 0.009	0.412 ± 0.010
A_T	0.725 ± 0.008		0.992 ± 0.001
A	0.543 ± 0.010	0.967 ± 0.003	

Spearman correlation coefficients for data on W1 hens below the diagonal and for data on WB hens above the diagonal

S.e.'s were obtained with R v3.1.1 and very slightly deviated from s.e.'s reported in GSE

To gain more insight, we calculated the loss in response to selection that occurs when applying a traditional model to individual or pooled data instead of a direct-indirect model to individual data. When applying a traditional model on individual data, the loss in total response to selection was 46.9% for W1 (Figure 4.1A) and 54.9% for WB (Figure 4.1C). When applying a traditional model on pooled data, the loss in total response to selection was 3.3% for W1 (Figure 4.1B) and 0.3% for WB (Figure 4.1D).

In conclusion, the loss in total response to selection will be large when using a traditional animal model on individual data, but will be small when using a traditional animal model on pooled data. However, this outcome may be specific to this dataset. Survival time in purebred laying hens was recorded in cages with four unrelated birds. Both direct and indirect genetic effects strongly influenced the trait. Group size, group composition, and the relative impact of direct and indirect genetic effects might influence the loss in total response to selection. For example, for body weight at 19 and 27 weeks of age, indirect genetic effects are expected to be small. In that case, an animal's A_T is mainly expressed in the phenotype of the animal itself. Consequently, we expect that more accurate estimated breeding values can be obtained when using individual instead of pooled data. Biscarini *et al.* (2008) found a correlation of ~ 0.75 between the estimated breeding values based on individual and pooled data, resulting in a large loss in response to selection when using pooled instead of individual data. Thus, using pooled data does not always seem to be a proper alternative and requires further research.

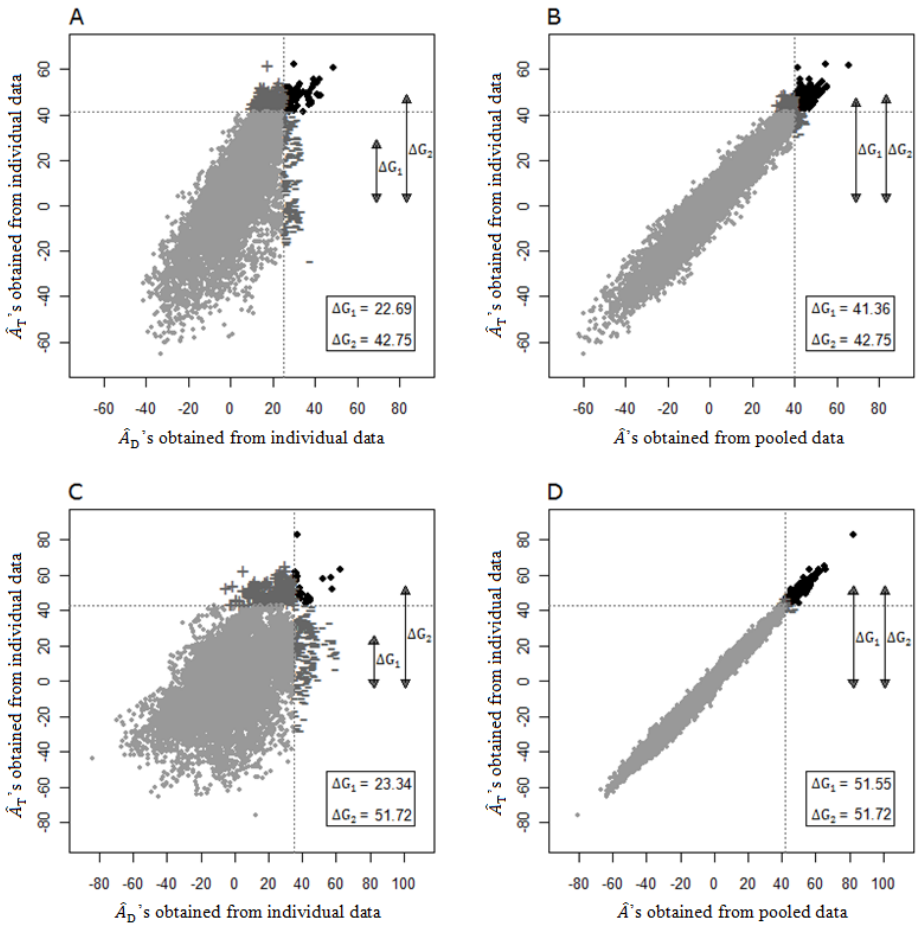

Figure 4.1 $\widehat{A}_T$'s obtained from individual data plotted against $\widehat{A}_D$'s obtained from individual data and $\widehat{A}$'s obtained from pooled data on survival time in laying hens. A and B for data on W1 hens. C and D for data on WB hens; ΔG_1 represents the total response to selection when selecting animals based on their $\widehat{A}_D$ obtained from individual data or $\widehat{A}$ obtained from pooled data; ΔG_2 represents the total response to selection when selecting animals based on their $\widehat{A}_T$ obtained from individual data.

4.4 Conclusions

Using pooled data, the total genetic variance and breeding values can be estimated, but the underlying direct and indirect genetic (co)variances and breeding values cannot. The most accurate estimates are obtained when group members belong to the same family. While quantifying the direct and indirect genetic effects is interesting from a biological perspective, obtaining the total

genetic effect is most important from an animal breeding perspective. When it is too difficult or expensive to obtain individual data, pooled data can be used to improve traits.

4.5 Acknowledgements

This research was financially supported by the Netherlands Organization for Scientific Research (NWO) and coordinated by the Dutch Technology Foundation (STW). The Institut de Sélection Animale B.V., a Hendrix Genetics Company, provided the data and was closely involved in this research. The authors thank Ewa Sell-Kubiak, Naomi Duijvesteijn and Sophie Eaglen for their valuable input.

4.6 Appendix A

This section demonstrates why direct and indirect (co)variances can be estimated from individual data, but cannot be estimated from pooled data.

Consider a situation where four base parents produce six offspring. Animals are kept in groups of two and individual phenotypes are recorded on all six offspring (Table 4.8).

Table 4.8 Example pedigree structure and group composition

Animal	Sire	Dam	Phenotype	Group
1	-	-	-	-
2	-	-	-	-
3	-	-	-	-
4	-	-	-	-
5	1	3	✓	1
6	2	4	✓	1
7	1	4	✓	2
8	2	3	✓	2
9	2	3	✓	3
10	2	4	✓	3

When analysing individual data with a direct-indirect animal model, the **Z**-matrices would be:

$$\mathbf{Z}_{\mathrm{D}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{Z}_{\mathrm{I}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

$\mathbf{Z}_{\mathrm{D}}$ and $\mathbf{Z}_{\mathrm{I}}$ are not identical, indicating that the direct and indirect genetic effects are estimated based on different information sources, enabling the model to distinguish between these two effects.

When analysing pooled data with a direct-indirect animal model, the **Z**-matrices would be:

$$\mathbf{Z}_{\mathrm{D}}^{*} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix},$$

$$\mathbf{Z}_{\mathrm{I}}^{*} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

$\mathbf{Z}_{\mathrm{D}}^{*}$ and $\mathbf{Z}_{\mathrm{I}}^{*}$ are identical, indicating that the direct and indirect genetic effects are estimated based on the same information source, causing complete confounding between direct and indirect genetic effects. The model will not be able to distinguish between these two effects.

4.7 Appendix B

Components of variance are determined by analysis of variance, where the full variance (σ_{z}^2) is partitioned into a between- (σ_{b}^2) and within-family component (σ_{w}^2). In this section, the derivation of σ_{z}^2, σ_{b}^2 and σ_{w}^2 are presented for three group compositions.

(i) When the group is composed of only one family, the A_T of a family is expressed n times in the same pooled record. Therefore, the record of interest is P^*/n.

$$\sigma_z^2 = \frac{\sigma_{P^*}^2}{n^2} = \frac{n\left(\sigma_{P_D}^2 + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma_{P_I}^2\right)}{n^2} + \frac{n(n-1)r\left(\sigma_{A_D}^2 + 2(n-1)\sigma_{A_{DI}} + (n-1)^2\sigma_{A_I}^2\right)}{n^2} = \frac{\sigma_{P_D}^2 + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma_{P_I}^2 + (n-1)r\sigma_{A_T}^2}{n}$$

$$\sigma_b^2 = r\sigma_{A_T}^2$$

$$\sigma_w^2 = \frac{\sigma_{P_D}^2 + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma_{P_I}^2 + (n-1)r\sigma_{A_T}^2}{n} - r\sigma_{A_T}^2$$

(ii) When the group is composed of two families, the A_T of a family is expressed $n/2$ times in the same pooled record. Therefore, the record of interest is $2P^*/n$.

$$\sigma_z^2 = \frac{4\sigma_{P^*}^2}{n^2} = \frac{4n\left(\sigma_{P_D}^2 + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma_{P_I}^2\right)}{n^2} + \frac{4n\left(\frac{n}{2}-1\right)r\left(\sigma_{A_D}^2 + 2(n-1)\sigma_{A_{DI}} + (n-1)^2\sigma_{A_I}^2\right)}{n^2} = \frac{4}{n}\left(\sigma_{P_D}^2 + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma_{P_I}^2 + \left(\frac{n}{2}-1\right)r\sigma_{A_T}^2\right)$$

$$\sigma_b^2 = r\sigma_{A_T}^2$$

$$\sigma_w^2 = \frac{4}{n}\left(\sigma_{P_D}^2 + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma_{P_I}^2 + \left(\frac{n}{2}-1\right)r\sigma_{A_T}^2\right) - r\sigma_{A_T}^2$$

(iii) When the group composition is random, the A_T of a family is only expressed once per pooled record. Therefore, the record of interest is P^*.

$$\sigma_z^2 = \sigma_{P^*}^2 = n\left(\sigma_{P_D}^2 + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma_{P_I}^2\right)$$

$$\sigma_b^2 = r\sigma_{A_T}^2$$

$$\sigma_w^2 = n\left(\sigma_{P_D}^2 + 2(n-1)\sigma_{P_{DI}} + (n-1)^2\sigma_{P_I}^2\right) - r\sigma_{A_T}^2$$

5

Bivariate analysis of individual and pooled data on social interaction traits: application to survival time and early egg production in laying hens

K. Peeters [1,2], J. Visscher[3], P. Bijma[1]

[1] Animal Breeding and Genomics Centre, Wageningen University, 6700 AH Wageningen, The Netherlands; [2] Research and Technology Centre, Hendrix Genetics, 5831 CK Boxmeer, The Netherlands; [3]Institut de Sélection Animale B.V., Hendrix Genetics, 5830 AC Boxmeer, The Netherlands

Submitted

Abstract

Individual data collection on group-housed animals can be difficult and expensive. Alternatively, data can be collected at group level, *e.g.* egg production or feed intake. Thus, phenotypic records on one animal can come from different levels. Bivariate analysis can introduce an issue with the non-genetic correlation between both traits, particularly for social interaction traits. For individual data, direct and indirect environmental effects are captured by two model terms: the group effect and the residual. However, for pooled data, direct and indirect environmental effects are captured by the residual only. Most statistical software programs cannot fit a correlation between the random group effect of the individual trait and residual of the pooled trait. This can result in biased genetic parameter estimates. We propose to adjust the model by adding a random group effect for the pooled trait and correlating it with the random group effect of the individual trait. The variance of the random group effect of the pooled trait needs to be set to a fixed value to avoid over-parameterization. A simulation study was conducted to validate whether the adjusted model indeed yields unbiased genetic parameter estimates. Subsequently, the genetic correlation between survival time (individual records) and early egg production (pooled records) in crossbred laying hens was estimated. The simulation study showed that the adjusted model, unlike the unadjusted model, resulted in unbiased genetic parameter estimates. With the adjusted model, the estimated genetic correlation between survival time and early egg production was negative (-0.09), but not significantly different from zero. In conclusion, to obtain unbiased genetic parameter estimates, it is necessary to add a random group effect for the pooled trait, to fix its variance, and to include a correlation with the random group effect for the individual trait.

Key words: Social interactions, indirect genetic effects, pooled data, survival time, early egg production, laying hens

5.1 Background

Individual data collection on group-housed animals can be difficult and expensive. Alternatively, data can be collected at group level, *e.g.* egg production or feed intake. Thus, phenotypic records on one animal can come from different levels. Bivariate analysis of individual and pooled data requires full understanding of the underlying genetic and environmental relationships. This becomes more complex when animals affect one another's phenotype through social interactions, which is common in group-housed animals. With social interactions, an individual has a direct (genetic) effect on its own phenotype and an indirect (genetic) effect on the phenotypes of its group mates (Griffing, 1967). Both direct and indirect genetic effects determine a population's potential to respond to selection, *i.e.* the total genetic variance (Bijma, 2011a). Previous studies have shown that direct, indirect and total genetic variances can be estimated from individual data, whereas only the total genetic variance can be estimated from pooled data (Muir, 2005; Bijma *et al.*, 2007; Peeters *et al.*, 2013). Here, we present a model that yields unbiased genetic parameter estimates for a bivariate analysis of individual and pooled data on social interaction traits. The model was validated through simulation and was then used to estimate genetic correlations between survival time (individual records) and early egg production (pooled records) in crossbred laying hens. This correlation was of interest because previous research suggested that selection for increased survival time delays the onset of lay, and the other way around (Lowry and Abplanalp, 1972; Craig *et al.*, 1975; Bhagwat and Craig, 1977; Ellen, 2008).

5.2 Methods

This study was performed in three steps. First, the bivariate analysis of individual and pooled data on social interaction traits was theoretically evaluated. We will show that, by using the default model, the analysis can result in biased genetic parameter estimates. We then propose a simple adjustment to the model to avoid this bias. Second, a simulation study was conducted to validate whether the adjusted model indeed yields unbiased genetic parameter estimates. Finally, the genetic correlation between survival time (ST; individual records) and early egg production (EEP; pooled records) in crossbred laying hens was estimated.

5.2.1 Theory

This section shows that, by using the default model, the bivariate analysis of individual and pooled data on social interaction traits can result in biased genetic parameter estimates, and introduces a simple adjustment to the model to avoid this bias.

Individual records

With social interactions, an individual record (P) consists of the direct genetic (A_D) and environmental (E_D) effect of the individual itself (i), and the indirect genetic (A_I) and environmental (E_I) effect of each of its $n-1$ group mates (j):

$$P_i = A_{D_i} + E_{D_i} + \sum_{i \neq j}^{n-1} A_{I_j} + \sum_{i \neq j}^{n-1} E_{I_j}, \quad [5.1]$$

where n is the number of individuals per group (Griffing, 1967). When estimating genetic parameters from individual data, the following direct-indirect animal model is commonly used (Bergsma *et al.*, 2008; Ellen *et al.*, 2008; Peeters *et al.*, 2012):

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_D a_D} + \mathbf{Z_I a_I} + \mathbf{Vg} + \mathbf{e}, \quad [5.2]$$

where $\mathbf{y}$ is a vector of individual records, $\mathbf{X}$ is an incidence matrix that links the individual records to fixed effects; $\mathbf{b}$ is a vector of fixed effects; $\mathbf{Z_D}$ is an incidence matrix that links the individual records to A_D's; $\mathbf{a_D}$ is a vector of A_D's; $\mathbf{Z_I}$ is an incidence matrix that links the individual records to A_I's; $\mathbf{a_I}$ is a vector of A_I's; $\mathbf{V}$ is an incidence matrix that links the individual records to random group effects; $\mathbf{g}$ is a vector of random group effects; and $\mathbf{e}$ is a vector of residuals. The random group effect σ^2_G accounts for the non-genetic covariance among group members (Bergsma *et al.*, 2008), *i.e.*:

$$\sigma^2_G = 2\sigma_{E_{DI}} + (n-2)\sigma^2_{E_I}. \quad [5.3]$$

The residual accounts for the remaining non-genetic variance (Bijma, 2011b):

$$\sigma^2_E = \sigma^2_{E_D} - 2\sigma_{E_{DI}} + \sigma^2_{E_I}. \quad [5.4]$$

Pooled records

A pooled record (P^*) is a summation of the individual records of all (k) group members:

$$P^* = \sum_{k=1}^{n} P_k. \quad [5.5]$$

It follows from Equation 5.1 and 5.5 that, with social interactions, a pooled record consists of the A_D and E_D of each group member, as well as their A_I and E_I that are each expressed $n-1$ times, *i.e.*:

$$P^* = \sum_{k=1}^{n} [A_{D_k} + E_{D_k} + (n-1)(A_{I_k} + E_{I_k})]. \quad [5.6]$$

Peeters *et al.* (2013) showed that direct and indirect genetic effects cannot be estimated from pooled data. However, a pooled record can be rewritten in terms of the total genetic effects (A_T) of all k group members. From an animal breeding perspective, A_T is of interest, because it is the genetic variance relevant for selection (Bijma, 2011a). An animal's A_T consists of a direct and indirect component,

$$A_{T_i} = A_{D_i} + (n-1)A_{I_i}, \quad [5.7]$$

where A_D is expressed in the phenotype of the animal itself and A_I is expressed in the phenotype of each group mate. Therefore, Equation 5.6 can be rewritten as:

$$P^* = \sum_{k=1}^{n} [A_{T_k} + E_{D_k} + (n-1)E_{I_k}]. \quad [5.8]$$

When estimating genetic parameters from pooled data, the following pooled animal model can be used (Peeters *et al.*, 2013):

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b}^* + \mathbf{Z}_\mathbf{T}^*\mathbf{a}_T + \mathbf{e}^*, \quad [5.9]$$

where $\mathbf{y}^*$ is a vector of pooled records, $\mathbf{X}^*$ is an incidence matrix linking the pooled records to fixed effects; $\mathbf{b}^*$ is a vector of fixed effects; $\mathbf{Z}_\mathbf{T}^*$ is an incidence matrix linking the pooled records to A_T's (each pooled record is linked to the A_T's of the k group members); $\mathbf{a}_\mathbf{T}$ is a vector of A_T's; and $\mathbf{e}^*$ is a vector of residuals. The residual accounts for the direct-indirect non-genetic (co)variances (Peeters *et al.*, 2013), *i.e.*:

$$\sigma^2_{E^*} = n\left[\sigma^2_{E_D} + 2(n-1)\,\sigma_{E_{DI}} + (n-1)^2\,\sigma^2_{E_I}\right]. \quad [5.10]$$

Bivariate analysis of individual and pooled data on social interaction traits can introduce an issue with the non-genetic correlation between both traits. For individual data, direct and indirect environmental effects are captured by two model terms, *i.e.* **g** and **e**. However, for pooled data, direct and indirect environmental effects are captured by one model term only, *i.e.* **e***. In most statistical software programs, correlations can only be fitted between genetic terms and between corresponding environmental terms, *e.g.* a correlation can be fitted between the residual of the individual and pooled trait. On the contrary, a correlation cannot be fitted between the random group effect of the individual trait and the residual of the pooled trait. This can bias the (genetic) parameter estimates. To overcome this problem, we propose to add a random group effect for

the pooled trait, which can then be correlated to the random group effect of the individual trait. However, the random group effect and residual of the pooled trait will be completely confounded, because there is only a single record per group. To avoid over-parameterization, the group variance of the pooled trait needs to be set to a fixed value, which makes the residual variance identifiable. The group variance should be set to a realistic value to avoid any of the other non-genetic parameters to be outside of their parameter space. We propose to use half of the pooled residual variance that was found in the univariate analysis of the pooled trait.
From this point onwards, a bivariate model without a random group effect for the pooled trait will be referred to as the unadjusted model, whereas the bivariate model with a random group effect for the pooled trait will be referred to as the adjusted model.

5.2.2. Simulation

To validate the theoretical expectations, a simulation study was conducted. Using R v2.12.2 (Venables *et al.*, 2011), a base population of 500 sires and 500 dams was simulated. Each animal in the base population was assigned direct and indirect breeding values for two social interaction traits. Breeding values were drawn from a multivariate normal distribution, where the direct and indirect genetic variances of both traits were set to 1.00 and all covariances were set to 0.25. Each sire was mated to a single randomly chosen dam, resulting in 12 offspring per mating and 6 000 offspring in total. Offspring breeding values were obtained (separately for $\mathrm{A}_{1_\mathrm{D}}$, $\mathrm{A}_{1_\mathrm{I}}$, $\mathrm{A}_{2_\mathrm{D}}$ and $\mathrm{A}_{2_\mathrm{I}}$) as $\mathrm{A} = \frac{1}{2}\mathrm{A}_{\mathrm{Sire}} + \frac{1}{2}\mathrm{A}_{\mathrm{Dam}} + \mathrm{MS}$. The direct and indirect Mendelian sampling terms were drawn from a multivariate normal distribution, where the direct and indirect genetic variances of both traits were set to 0.50 and all covariances were set to 0.125. Each offspring was also assigned direct and indirect environmental values for each trait. The environmental values were drawn from a multivariate normal distribution where the direct and indirect environmental variances of both traits were set to 2.00 and all covariances were set to 0.50. Four animals were randomly assigned to a group, resulting in 1 500 groups. For Trait 1, individual phenotypes were obtained by adding up the direct and indirect genetic and environmental components according to Equation 5.1. For Trait 2, pooled records were obtained by adding up the direct and indirect genetic and environmental components according to Equation 5.6. One hundred replicates were produced.

An unadjusted bivariate model was fitted in ASReml v3.0 (Gilmour *et al.*, 2009), where the individual records were evaluated with a direct-indirect animal model (Equation 5.2) and the pooled records were evaluated with a pooled animal model (Equation 5.9). The genetic (co)variance structure was:

$$\mathrm{Var}\begin{bmatrix}\mathbf{a}_{1\_D}\\ \mathbf{a}_{1\_I}\\ \mathbf{a}_{2\_T}\end{bmatrix} = \begin{bmatrix}\sigma^2_{A_{1\_D}} & \sigma_{A_{1\_DI}} & \sigma_{A_{1\_D\_2\_T}}\\ \sigma_{A_{1\_DI}} & \sigma^2_{A_{1\_I}} & \sigma_{A_{1\_I\_2\_T}}\\ \sigma_{A_{1\_D\_2\_T}} & \sigma_{A_{1\_I\_2\_T}} & \sigma^2_{A_{2\_T}}\end{bmatrix} \otimes \mathbf{A}.$$

The residual (co)variance structure was:

$$\mathrm{Var}\begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}^*_2\end{bmatrix} = \begin{bmatrix}\sigma^2_{E_1} & \sigma_{E^{\times}_{12}}\\ \sigma_{E^{\times}_{12}} & \sigma^2_{E^*_2}\end{bmatrix} \otimes \mathbf{I}.$$

Based on the theory, we expect a bias in the genetic parameter estimates because the correlation between the random group effect of the individual trait and the residual of the pooled trait is ignored. Therefore, an adjusted bivariate model was fitted in ASReml v3.0, where the pooled records were evaluated with a pooled animal model (Equation 5.9) that included a random group effect. This random group effect was correlated to the random group effect of the individual trait. The (co)variance structure of the random group effect was:

$$\mathrm{Var}\begin{bmatrix}\mathbf{g}_1\\ \mathbf{g}_2\end{bmatrix} = \begin{bmatrix}\sigma^2_{G_1} & \sigma_{G_{12}}\\ \sigma_{G_{12}} & \sigma^2_{G_2}\end{bmatrix} \otimes \mathbf{I},$$

where $\sigma^2_{G_2}$ was fixed as it is completely confounded with $\sigma^2_{E^*_2}$. The $\sigma^2_{G_2}$ was fixed at a value of 46, which is half of the expected pooled residual variance (Equation 5.10).

5.2.3 Survival time and early egg production

To investigate whether selection for increased survival time in laying hens will indeed delay the onset of lay, we estimated the genetic correlation between survival time (ST) and early egg production (EEP) using the adjusted model.

The dataset was part of the pre-existing database of Hendrix Genetics (The Netherlands) and contained routinely collected data for breeding value estimation. Animal Care and Use Committee approval was therefore not required.

Two commercial purebred White Leghorn layer lines, W1 and WB, produced 15 012 crossbred laying hens (*Gallus gallus*) of which 7 668 were W1xWB (♂x♀) and 7 344 were WBxW1 (♂x♀). On average, each cross was produced by mating ~50 sires to ~705 randomly assigned dams, where dams were nested within sires.

At approximately 17 weeks of age, the hens were placed in two laying houses. The laying houses consisted of four or five double rows, and each row consisted of three levels. Interaction with back neighbours was possible, but interaction with side neighbours was prevented. Four hens of the same cross were randomly assigned to a cage. Thus, other than by chance, cage mates were unrelated. Hens were not beak-trimmed. Further details on housing conditions and management are in Peeters *et al.* (2012).

The traits of interest were survival time and early egg production. Survival time (ST) was defined as the number of days from the moment the animal was placed in the laying house until either death or the end of the experiment, with a maximum of 398 days. Early egg production (EEP) was defined as the number of eggs produced by four birds within a cage, from the age of 129 days until 143 days (two-week period). Because EEP is a pooled record, all four birds had to be alive during this two-week period. Therefore, EEP was set to missing for 983 cages.

ST was analysed both as an individual and pooled trait. This was done for two reasons. First, it gives the opportunity to compare the variance components for ST obtained from individual and pooled records. Second, it gives the opportunity to compare the correlations between (i) pooled ST and pooled EEP, (ii) individual ST and pooled EEP from the unadjusted model, and (iii) individual ST and pooled EEP from the adjusted model. Pooled ST records were obtained by summing the individual ST records within the group (Equation 5.5). When pooling data, all four records in a group had to be known. Therefore, 23 cages were excluded from the analysis, since some individual records were missing.

The dataset was split up to enable separate analyses for W1xWB and WBxW1 data, as a previous analysis showed that reciprocal effects influence ST (Peeters *et al.*, 2012)

Univariate analyses for survival time and early egg production

When using individual records on ST, direct and indirect genetic parameters were estimated by fitting a direct-indirect animal model in ASReml v3.0 (Equation 5.2). The fixed effects were an interaction term for each laying house by row by level combination, a fixed effect for the content of the back cage (empty or contains hens) and a covariate for the average number of survival days in the back cage.

When using pooled records on ST and EEP, total genetic parameters were estimated by fitting a pooled animal model in ASReml v3.0 (Equation 5.9). For ST, the same fixed effects were used as in the previous model; for EEP, only interaction term for each laying house by row by level combination was used.

Additional analysis on EEP was performed to check for potential bias due to data pre-selection, as 26% of the EEP records were non-randomly set to missing. Whether or not a record was set to missing depended on the mortality within a cage in the first two weeks of egg production. This type of pre-selection can cause potential bias in the genetic parameter estimation (Henderson, 1975). To check for potential bias due to pre-selection, a bivariate analysis was fitted between EEP and the pre-selection parameter (Arnason, 1999; Ducro, 2010; Albertsdóttir *et al.*, 2011). The pre-selection parameter was a binary 0-1 trait, where all cages with a missing record were assigned '0', whereas cages with an EEP record were assigned '1'. The necessity of a pre-selection parameter will be judged by the change in total genetic variance for EEP when adding the pre-selection parameter as a second trait to the model.

Genetic correlations between survival time and early egg production

A bivariate model was fitted in ASReml v3.0, where both ST and EEP are pooled records. There is no issue with the non-genetic correlation in a bivariate analysis of two pooled traits and, consequently, there is no bias in genetic parameter estimates. The genetic (co)variance structure was:

$$\mathrm{Var}\begin{bmatrix}\mathbf{a}_{\mathrm{ST_T}}\\ \mathbf{a}_{\mathrm{EEP_T}}\end{bmatrix} = \begin{bmatrix}\sigma^2_{\mathrm{A_{ST_T}}} & \sigma_{\mathrm{A_{ST\ EEP_T}}}\\ \sigma_{\mathrm{A_{ST\ EEP_T}}} & \sigma^2_{\mathrm{A_{EEP_T}}}\end{bmatrix} \otimes \mathbf{A}.$$

The parameter of interest was the genetic correlation between total breeding values for ST and EEP ($\mathrm{r}_{\mathrm{A_{ST\ EEP_T}}}$). This unbiased estimate can be used as a reference when comparing it to correlations that are obtained from the unadjusted and adjusted bivariate model, where ST is an individual record and EEP is a pooled record.

The genetic (co)variance structure for the unadjusted and adjusted model was:

$$\mathrm{Var}\begin{bmatrix}\mathbf{a}_{\mathrm{ST_D}}\\ \mathbf{a}_{\mathrm{ST_I}}\\ \mathbf{a}_{\mathrm{EEP_T}}\end{bmatrix} = \begin{bmatrix}\sigma^2_{\mathrm{A_{ST_D}}} & \sigma_{\mathrm{A_{ST_{DI}}}} & \sigma_{\mathrm{A_{ST_D\ EEP_T}}}\\ \sigma_{\mathrm{A_{ST_{DI}}}} & \sigma^2_{\mathrm{A_{ST_I}}} & \sigma_{\mathrm{A_{ST_I\ EEP_T}}}\\ \sigma_{\mathrm{A_{ST_D\ EEP_T}}} & \sigma_{\mathrm{A_{ST_I\ EEP_T}}} & \sigma^2_{\mathrm{A_{EEP_T}}}\end{bmatrix} \otimes \mathbf{A}.$$

The parameters of interest were the direct-total genetic correlation ($\mathrm{r}_{\mathrm{A_{ST_D EEP_T}}}$) and the indirect-total genetic correlation ($\mathrm{r}_{\mathrm{A_{ST_I EEP_T}}}$) between ST and EEP. Based on the genetic (co)variances, $\mathrm{r}_{\mathrm{A_{ST\ EEP_T}}}$ was calculated (See *Appendix*).

5.3 Results and discussion

5.3.1 Simulation

Results show that for the unadjusted bivariate model most genetic parameter estimates were biased (Table 5.1). In particular $\sigma_{A_{1\_D\_2\_T}}$ and $\sigma_{A_{1\_I\_2\_T}}$ showed large deviations. On the contrary, for the adjusted bivariate model, none of the genetic parameter estimates were biased (Table 5.1). This validates that to obtain unbiased genetic parameter estimates, a random group effect for the pooled trait needs to be added and correlated to the random group effect of the individual trait. The variance of the random group effect of the pooled trait needs to be set to a fixed value to avoid over-parameterization. The variance of the random group effect of the pooled trait was fixed at 46, which is half of the expected pooled residual variance (Equation 5.10). However, a 25% increase or decrease in the variance of the random group effect for the pooled trait yielded the exact same genetic parameter estimates.

Table 5.1 True and estimated genetic parameters with s.e. obtained from an unadjusted and adjusted bivariate model

	True parameters	Estimated parameters Unadjusted model	Estimated parameters Adjusted model
$\sigma^2_{A_{1\_D}}$	1.00	1.37 ± 0.20	1.01 ± 0.18
$\sigma^2_{A_{1\_I}}$	1.00	1.45 ± 0.17	1.00 ± 0.15
$\sigma^2_{A_{2\_T}}$	11.50	17.70 ± 2.26	11.64 ± 2.11
$\sigma_{A_{1\_DI}}$	0.25	0.41 ± 0.63	0.26 ± 0.13
$\sigma_{A_{1\_D\_2\_T}}$†	1.00	2.99 ± 0.41	1.02 ± 0.46
$\sigma_{A_{1\_I\_2\_T}}$††	1.00	3.24 ± 0.32	1.02 ± 0.42

† $\sigma_{A_{1\_D\_2\_T}} = \mathrm{Covar}(A_{1\_D}; A_{2\_D} + (n-1)A_{2\_I}) = \sigma_{A_{12\_D}} + (n-1)\,\sigma_{A_{1\_D\_2\_I}}$

†† $\sigma_{A_{1\_I\_2\_T}} = \mathrm{Covar}(A_{1\_I}; A_{2\_D} + (n-1)A_{2\_I}) = \sigma_{A_{1\_I\_2\_D}} + (n-1)\,\sigma_{A_{12\_I}}$

5.3.2 Survival time and early egg production

Univariate analysis of survival time

Table 5.2 shows the estimated variance components for ST using either individual or pooled records. The genetic parameters for individual ST were previously estimated for the same population, but on a slightly larger dataset by Peeters *et al.*

(2012). They are re-estimated to facilitate the comparison with the genetic parameter estimates for pooled ST.

The direct-indirect animal model on individual ST yields estimates of $\sigma^2_{A_D}$, $\sigma_{A_{DI}}$ and $\sigma^2_{A_I}$. Based on these components, $\sigma^2_{A_T}$ was calculated, as $\sigma^2_{A_T} = \sigma^2_{A_D} + 2(n-1)\,\sigma_{A_{DI}} + (n-1)^2\sigma^2_{A_I}$ (Bijma, 2007). The pooled animal model on pooled ST yields estimates of $\sigma^2_{A_T}$ directly (Peeters *et al.*, 2013). For W1xWB, the estimates of $\sigma^2_{A_T}$ from individual and pooled records are very similar. Similarly, Peeters *et al.* (2013) found that, for W1 and WB, the estimates of $\sigma^2_{A_T}$ from individual and pooled survival data were very similar. However, for WBxW1, the estimated $\sigma^2_{A_T}$ from individual data was 1.26 times larger than the estimate from pooled data (Table 5.2). The direct-indirect animal model on individual data also yielded estimates of σ^2_G and σ^2_E. Peeters *et al.* (2013) showed that based on these components, the expected $\sigma^2_{E^*}$ can be calculated, as $\sigma^2_{E^*} = n^2\sigma^2_G + n\sigma^2_E$. For both crosses, the expected $\sigma^2_{E^*}$ calculated based on individual data and estimate of $\sigma^2_{E^*}$ from pooled data are very similar (Table 5.2).

Table 5.2 Estimated parameters with s.e. for survival time using individual or pooled data

	W1xWB	WBxW1
Individual data		
$\sigma^2_{A_D}$	615 ± 170	935 ± 222
$\sigma_{A_{DI}}$	-182 ± 96	-715 ± 140
$\sigma^2_{A_I}$	509 ± 106	776 ± 150
$\sigma^2_{A_T}$†	4 105 ± 1 055	3 627 ± 1 088
σ^2_G	2 014 ± 263	2 350 ± 310
σ^2_E	11 722 ± 301	15 728 ± 460
Expected $\sigma^2_{E^*}$††	79 108 ± 4 061	100 506 ± 4 736
σ^2_P	15 878 ± 289	21 341 ± 402
Pooled data		
$\sigma^2_{A_T}$	4 079 ± 1 077	2 835 ± 1 008
$\sigma^2_{E^*}$	79 154 ± 4 116	103 201 ± 4 719
$\sigma^2_{P^*}$	95 470 ± 3 306	114 541 ± 3 913

†$\sigma^2_{A_T} = \sigma^2_{A_D} + 2(n-1)\,\sigma_{A_{DI}} + (n-1)^2\sigma^2_{A_I}$

††$\sigma^2_{E^*} = n^2\sigma^2_G + n\sigma^2_E$

Univariate analysis of early egg production

The total genetic variance explained 55% of the pooled phenotypic variance for W1xWB and 50% for WBxW1 ($T^2 = 4\sigma^2_{A_T}/\sigma^2_{P^*}$; Table 5.3). Bivariate analysis of EEP and the pre-selection parameter did not change $\sigma^2_{A_T}$ substantially (identical for W1xWB; from 4.90 to 4.88 for WBxW1). Therefore, the pre-selection parameter was omitted from further analyses.

Table 5.3 Estimated variance components with s.e. for early egg production using pooled data

	W1xWB	WBxW1
$\sigma^2_{A_T}$	4.92 ± 0.85	4.90 ± 0.94
$\sigma^2_{E^*}$	16.05 ± 2.29	19.94 ± 2.61
$\sigma^2_{P^*}$	35.72 ± 1.75	39.53 ± 1.95

In the literature, heritability estimates for individual EEP range from 0.26 to 0.43 (Besbes *et al.*, 1992; Anang *et al.*, 2000; Nurgiartiningsih *et al.*, 2004; Wolc *et al.*, 2007). Because these estimates were obtained from data on individually housed birds, they are estimates of direct genetic variance, rather than total genetic variance. In addition, GxE (genotype by environment) effects might occur (individual *vs* group housing). Moreover, the error variances of individual and pooled data are not necessarily proportional (as shown by Peeters *et al.* (2013) and in the footnote of Table 5.2). Nurgiartiningsih *et al.* (2004) and Biscarini *et al.* (2010b) also reported estimates obtained from pooled data on group-housed birds, which range from 0.30 to 0.38. Overall, in the literature, genetics explained a smaller proportion of the phenotypic variance than found in this study. However, the EEP periods were longer than in this study (4 to 7 weeks *vs* 2 weeks). Because egg production has a lower heritability in later stages of production (Besbes *et al.*, 1992; Anang *et al.*, 2000; Nurgiartiningsih *et al.*, 2004; Wolc *et al.*, 2007), an extension of the EEP period will result in a decreased heritability.

Bivariate analysis of survival time and early egg production

Table 5.4 shows estimated genetic correlations between (i) pooled ST and pooled EEP, (ii) individual ST and pooled EEP from the unadjusted model, and (iii) individual ST and pooled EEP from the adjusted model. For analysis (i), $r_{A_{ST\,EEP_T}}$ was obtained directly. For analysis (ii) and (iii), $r_{A_{ST\,EEP_T}}$ was calculated from the underlying estimated genetic (co)variances (See *Appendix*). The $r_{A_{ST\,EEP_T}}$ estimated from the adjusted model and from the pooled analysis were very similar, whereas the estimate from the unadjusted model had the opposite sign. This supports the above findings that the adjusted model is superior over the unadjusted model.

Table 5.4 Genetic correlations, with s.e., between survival time (ST) and early egg production (EEP)

	W1xWB	WBxW1
ST and EEP pooled		
$r_{A_{ST\,EEP_T}}$	-0.09 ± 0.15	-0.15 ± 0.19
ST individual and EEP pooled (Unadjusted model)		
$r_{A_{ST_D EEP_T}}$	-0.05 ± 0.15	0.11 ± 0.15
$r_{A_{ST_I EEP_T}}$	0.17 ± 0.10	0.04 ± 0.11
$r_{A_{ST\,EEP_T}}$	0.17 ± 0.10	0.11 ± 0.12
ST individual and EEP pooled (Adjusted model)		
$r_{A_{ST_D EEP_T}}$	-0.15 ± 0.16	0.12 ± 0.15
$r_{A_{ST_I EEP_T}}$	-0.03 ± 0.13	-0.11 ± 0.14
$r_{A_{ST\,EEP_T}}$	-0.09 ± 0.15	-0.09 ± 0.17

The estimated $r_{A_{ST\,EEP_T}}$ from both the pooled and the adjusted model was slightly negative but not significantly different from zero (Table 5.4). This is a weak indication of an unfavourable genetic correlation between ST and EEP. In the literature, multiple selection experiments indicate that survival time and early egg production in laying hens are indeed negatively correlated [Lowry and Abplanalp, 1972; Craig *et al.*, 1975; Bhagwat and Craig, 1977; Ellen, 2009]. Huges (1973) reported a biological link between survival time and early egg production, since the gonadal hormones that trigger the onset of lay also cause feather pecking and

cannibalism. The underlying correlations, $r_{A_{ST_D\,EEP_T}}$ and $r_{A_{ST_I\,EEP_T}}$, did not differ significantly from zero and showed no systematic pattern when both crosses were compared. In conclusion, our analysis provides a weak indication for a negative genetic correlation between survival time and early egg production.

In this study, we have focused on traits that are affected by social interactions. However, our findings are relevant whenever individual and pooled data on group-housed animals are analysed jointly. For the individually recorded trait, group effects may occur for other reasons than social interactions; *e.g.* differences in the physical environment among groups. When group effects are present, this may create a covariance between the random group effect of the individual trait and the residual of the pooled trait, which may bias genetic parameter estimates when using an unadjusted model. We have shown that this issue can be solved by adding a random group effect with a fixed variance to the model for the pooled trait, and allowing for a covariance with the random group effect for the individual trait.

5.4 Conclusion

To obtain unbiased genetic parameter estimates, it is necessary to add a random group effect for the pooled trait, to fix its variance, and to include a correlation with the random group effect for the individual trait. Application to survival time and early egg production in laying hens yielded a slightly negative estimate of the genetic correlation between both traits, which was not significantly different from zero.

5.5 Acknowledgements

This research was financially supported by the Netherlands Organization for Scientific Research (NWO) and coordinated by the Dutch Technology Foundation (STW). Institut de Sélection Animale B.V., a Hendrix Genetics company, provided the data and was closely involved in this research. The authors thank Ewa Sell-Kubiak, Naomi Duijvesteijn and Sophie Eaglen for their valuable input.

5.6 Appendix

To calculate the total genetic correlation between two traits, $\sigma^2_{A_{1\_T}}$, $\sigma^2_{A_{2_T}}$and $\sigma_{A_{12\_T}}$ need to be known, as $r_{12\_T} = \sigma_{A_{12\_T}} \Big/ \sqrt{\sigma^2_{A_{1\_T}} \sigma^2_{A_{2\_T}}}$. A bivariate analysis of individual data (Trait 1) and pooled data (Trait 2) does not yield $\sigma^2_{A_{1\_T}}$ and $\sigma_{A_{12\_T}}$ directly. Based on the direct-indirect (co)variances from Trait 1, $\sigma^2_{A_{1\_T}}$ was calculated, as

$\sigma^2_{A_{1\_T}} = \mathrm{Var}\left(A_{1\_D} + (n-1)A_{1\_I}\right) = \sigma^2_{A_{1\_D}} + 2(n-1)\,\sigma_{A_{1\_DI}} + (n-1)^2\,\sigma^2_{A_{1\_I}}$.
Based on the direct-total and indirect-total covariance between traits, $\sigma_{A_{12\_T}}$ was calculated, as $\sigma_{A_{12\_T}} = \mathrm{Covar}\left(A_{1\_D} + (n-1)A_{1\_I}; A_{2\_T}\right) = \sigma_{A_{1\_D\_2\_T}} + (n-1)\,\sigma_{A_{1\_I\_2\_T}}$. Therefore, $r_{12\_T} = \frac{\sigma_{A_{1\_D\_2\_T}} + (n-1)\,\sigma_{A_{1\_I\_2\_T}}}{\sqrt{\left(\sigma^2_{A_{1\_D}} + 2(n-1)\,\sigma_{A_{1\_DI}} + (n-1)^2\,\sigma^2_{A_{1\_I}}\right)\sigma^2_{A_{2\_T}}}}$.

6

General discussion

6.1 Introduction

In this thesis, we estimated genetic parameters for survival time in non-beak-trimmed group-housed laying hens. The experimental set-up varied between chapters, either because of a difference in the level of data collection (individual data *vs* pooled data) and/or the within-group relatedness (unrelated *vs* half-sibs). The experimental set-up determined whether or not direct and indirect genetic parameters could be estimated. In more detail:

- In Chapter 2, we used individual survival time data on two purebred lines (W1 and WB) and their reciprocal cross (W1xWB and WBxW1), where unrelated hens of the same line or cross were kept in groups of four. This experimental set-up allowed us to estimate direct, indirect and total genetic variances. T^2 was 0.19 for W1, 0.16 for WB, 0.26 for W1xWB and 0.17 for WBxW1. We showed that survival time was strongly influenced by the social interactions among birds. For crossbreds, the indirect genetic variance was found to be approximately the same size as the direct genetic variance. Consequently, following Equation 1.5, the indirect genetic variance contributed nine times more to the total genetic variance than the direct genetic variance.
- In Chapter 3, we used individual survival time data on the crossbred offspring of three sire lines (W1, W5 and W6), where half-sibs from the same sire were kept in groups of five. No dam information was available. Therefore, we used a sire model instead of an animal model. This experimental set-up did not allow us to estimate the direct and indirect genetic (co)variances. However, the total genetic variance could be estimated. T^2 was 0.30 for W1, 0.16 for W5 and 0.12 for W6.
- In Chapter 4, we used pooled survival time data on two purebred lines (W1 and WB), where unrelated hens of the same line were kept in groups of four. This experimental set-up allowed us to estimate the total genetic variance, but not the underlying direct and indirect genetic (co)variances. T^2 was 0.15 for W1 and 0.13 for WB.
- In Chapter 5, we used pooled survival time data on a reciprocal cross (W1xWB and WBxW1), where unrelated hens of the same cross were kept in groups of four. This experimental set-up allowed us to estimate the total genetic variance, but not the underlying direct and indirect genetic (co)variances. T^2 was 0.17 for W1xWB and 0.10 for WBxW1.

In conclusion, throughout this thesis we showed that the level of data collection (individual data *vs* pooled data) and the within-group relatedness (unrelated *vs* half-sibs) can cause confounding of direct and indirect genetic effects. This prevented us from estimating direct and indirect genetic (co)variances for certain experimental set-ups in Chapter 3, 4 and 5. In Chapter 4 we briefly touched upon the fact that the experimental set-up also affects the phenotypic variance. Consequently, the experimental set-up affects T^2, as T^2 is the total heritable variance relative to phenotypic variance. This indicates that T^2, as a measure of inheritance, should be used with care. The same holds true for the classic h^2. It seems that most quantitative geneticists are not aware of the dependency of T^2 and h^2 on the experimental set-up. Therefore, this will be the first topic that I will discuss in my general discussion.

As a second topic of my general discussion, I will discuss the practical implementation and coinciding hurdles of genetically improving social interaction traits. How can we improve survival in non-beak-trimmed group-housed laying hens? What is the impact of improper modelling? How do we improve social interaction traits when animals are kept in large groups?

6.2 h^2 and T^2 as a measures of inheritance

For non-social interaction traits, the heritability (h^2) is used as a measure of inheritance, where the direct genetic variance ($\sigma^2_{A_D}$) is expressed relative to the phenotypic variance (σ^2_P);

$$h^2 = \frac{\sigma^2_{A_D}}{\sigma^2_P}.$$

Analogously, for social interaction traits, T^2 is used as a measure of inheritance, where the total genetic variance ($\sigma^2_{A_T}$) is expressed relative to σ^2_P;

$$T^2 = \frac{\sigma^2_{A_T}}{\sigma^2_P}.$$

Throughout this thesis it became clear that, for social interaction traits, the experimental set-up affects the phenotypic variance (σ^2_P). Consequently, the experimental set-up also affects T^2. The dependency of σ^2_P and T^2 on the experimental set-up has been briefly discussed by Bergsma *et al.* (2008) and in Chapter 4 of this thesis. However, the consequences of this dependency have not been discussed. Therefore, I will look at T^2 as a measure of inheritance for social interaction traits. In addition, h^2 also depends on the experimental set-up. Therefore, I will also look at h^2 as a measure of inheritance for non-social interaction traits.

Four experimental set-ups that differ in the level of data collection (individual *vs* pooled data) and/or the within-group relatedness (unrelated *vs* half-sibs) will be compared (Figure 6.1).

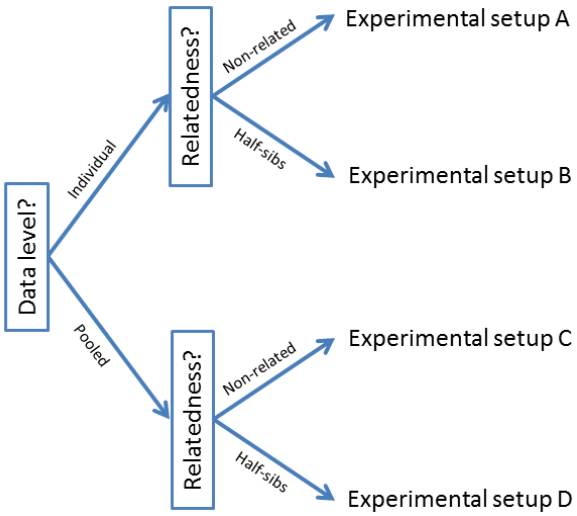

Figure 6.1 Overview of the four experimental set-ups that either differed in the level of data collection (individual *vs* pooled data) and/or the within-group relatedness (unrelated *vs* half-sibs)

First, T^2 and h^2 will be theoretically derived for each experimental set-up. This will demonstrate that, despite having the same underlying genetic and environmental parameters, both T^2 and h^2 can differ between experimental set-ups. In addition, for each experimental set-up, I will give an overview of the animal and sire models that can be used to estimate variance components (social interaction traits only; Box A, B, C and D). The aim is to gain insight into the underlying direct and indirect genetic (co)variances of the estimated variance components. This will help geneticists to understand the differences between experimental set-ups and models, which is essential for a correct interpretation of estimates from the literature.

Second, a simulation will be conducted to confirm the theoretical derivation of T^2 and the underlying direct and indirect genetic (co)variances of the estimated variance components as presented in Box A, B, C and D.

Third, empirical results from this thesis will be used to illustrate how T^2 for survival time will differ between experimental set-ups.
Finally, empirical results from literature will be evaluated. I will give examples of measures of inheritance that differ between experimental set-ups.
For correct biological inference and comparison of studies, it is important to realize that the experimental set-up affects σ_P^2 and T^2, but does not affect $\sigma_{A_T}^2$. The potential to respond to selection is identical for all four experimental set-ups.

6.2.1 Theory

T^2 as a measure of inheritance

The dependency of T^2 on the experimental set-up solely originates from the dependency of σ_P^2 on the experimental set-up. The following, therefore, focusses on the phenotypic variance.
Depending on the scenario, the following phenotypic variance was found:

- Experimental set-up A, individual data & unrelated birds

As shown in Equation 1.2:

$$P_i = A_{D_i} + \sum_{i \neq j}^{n-1} A_{I_j} + E_{D_i} + \sum_{i \neq j}^{n-1} E_{I_j}.$$

Consequently,

$$\sigma_P^2 = \sigma_{A_D}^2 + (n-1)\sigma_{A_I}^2 + \sigma_{E_D}^2 + (n-1)\sigma_{E_I}^2.$$

- Experimental set-up B, individual data & half-sibs

In Equation 1.2, the genetic components can be written in terms of the sire, dam and Mendelian sampling term. Because all animals within a cage have the same sire, $\sum_{i \neq j}^{n-1} A_{Sire_{I_j}}$ can be replaced by $(n-1)A_{Sire_{I_j}}$:

$$P_i = \frac{1}{2}\left[A_{Sire_{D_i}} + (n-1)A_{Sire_{I_j}}\right] + \frac{1}{2}\left[A_{Dam_{D_i}} + \sum_{i \neq j}^{n-1} A_{Dam_{I_j}}\right] + MS_{D_i} + \sum_{i \neq j}^{n-1} MS_{I_j}$$

$$+E_{D_i} + \sum_{i \neq j}^{n-1} E_{I_j}.$$

Based on the above equation and Equation 1.4, P_i can be simplified by rewriting the sire component in terms of the total genetic effect:

$$P_i = \frac{1}{2}A_{Sire_{T_i}} + \frac{1}{2}\left[A_{Dam_{D_i}} + \sum_{i \neq j}^{n-1} A_{Dam_{I_j}}\right] + MS_{D_i} + \sum_{i \neq j}^{n-1} MS_{I_j} + E_{D_i} + \sum_{i \neq j}^{n-1} E_{I_j}.$$

Consequently (as shown by Bergsma *et al.*, 2008),

$$\sigma_P^2 = \frac{1}{4}\sigma_{A_T}^2 + \frac{3}{4}\left[\sigma_{A_D}^2 + (n-1)\sigma_{A_I}^2\right] + \sigma_{E_D}^2 + (n-1)\sigma_{E_I}^2.$$

- Experimental set-up C, pooled data & unrelated birds

A pooled phenotype (P^*) is obtained by summing the individual phenotypes within a group:

$$P^* = \sum_{k=1}^{n} P_k$$

where k refers to the group members.

Based on the above equation and Equation 1.2 and 1.4, P^* can be rewritten in terms of total genetic effects, and direct and indirect environmental effects:

$$P^* = \sum_{k=1}^{n} \left[A_{T_k} + E_{D_k} + (n-1)E_{I_k}\right].$$

Consequently (as shown in Chapter 4),

$$\sigma^2_{P^*} = n\left[\sigma^2_{A_T} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right].$$

- Experimental set-up D, pooled data & half-sibs

As shown in the previous section:

$$P^* = \sum_{k=1}^{n} \left[A_{T_k} + E_{D_k} + (n-1)E_{I_k}\right].$$

The genetic component can be written in terms of the sire, dam and Mendelian sampling term. Because all animals within a cage have the same sire, $\sum_{k=1}^{n} A_{Sire_{T_k}}$ can be replaced by $nA_{Sire_{T_k}}$:

$$P^* = \frac{1}{2}nA_{Sire_{T_k}} + \sum_{k=1}^{n}\left[\frac{1}{2}A_{Dam_{T_k}} + MS_{T_k} + E_{D_k} + (n-1)E_{I_k}\right].$$

Consequently,

$$\sigma^2_{P^*} = n\left[\frac{1}{4}(n+3)\sigma^2_{A_T} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right].$$

Experimental set-ups C and D result in a pooled phenotypic variance ($\sigma^2_{P^*}$, phenotypic variance of the sum of the n individual phenotypes within a group) rather than an individual phenotypic variance (σ^2_P). Therefore, when calculating T^2, $\sigma^2_{P^*}$ was divided by the number of group members:

- Experimental set-up A, individual data & unrelated birds

$$T^2 = \sigma^2_{A_T} / \left[\sigma^2_{A_D} + (n-1)\sigma^2_{A_I} + \sigma^2_{E_D} + (n-1)\sigma^2_{E_I}\right]$$

- Experimental set-up B, individual data & half-sibs

$$T^2 = \sigma^2_{A_T} / \left[\frac{3}{4}\left[\sigma^2_{A_D} + (n-1)\sigma^2_{A_I}\right] + \frac{1}{4}\sigma^2_{A_T} + \sigma^2_{E_D} + (n-1)\sigma^2_{E_I}\right]$$

- Experimental set-up C, pooled data & unrelated birds

$$T^2 = \sigma^2_{A_T} / \left[\sigma^2_{A_T} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$$

- Experimental set-up D, pooled data & half-sibs

$$T^2 = \sigma^2_{A_T} / \left[\frac{1}{4}(n+3)\sigma^2_{A_T} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$$

These equations show that, in the majority of cases, a higher level of data collection (using pooled data instead of individual data) and increased within-group relatedness will result in increased phenotypic variance and, consequently, a lower T^2.

For survival time in laying hens, T^2 might drop up to 42% when using experimental set-up D compared to experimental set-up A (more details in *Section 6.2.3*.).

To facilitate comparison of estimates of T^2 in the scientific literature, one would ideally standardize T^2, so that estimates can be compared irrespective of the experimental set-up. Since the numerator of T^2, $\sigma^2_{A_T}$, is the same for all four experimental set-ups, standardizing the definition of T^2 implies standardizing the denominator. I propose to use phenotypic variance for pooled data on unrelated group members, divided by the number of group members, *i.e.* $\sigma^2_{A_T} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}$. This yields the following standardized definition of T^2:

$$T^2_{Standardized} = \sigma^2_{A_T} / \left[\sigma^2_{A_T} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right].$$

I propose to use the phenotypic variance for pooled data on unrelated group members rather than the phenotypic variance for individual data on unrelated group members because pooled phenotypic variance can be calculated based on estimates from individual data but not the other way around.

In practice, for individual data (analysed with an animal model) this would imply:

$$T^2_{Standardized} = \sigma^2_{A_T} / \left[\sigma^2_{A_T} + n\sigma^2_{Cage} + \sigma^2_E\right],$$

as $\sigma^2_{Cage} = 2\sigma_{E_{DI}} + (n-2)\sigma^2_{E_I}$ and $\sigma^2_E = \sigma^2_{E_D} - 2\sigma_{E_{DI}} + \sigma^2_{E_I}$

and for pooled data (analysed with an animal model) this would imply:

$$T^2_{Standardized} = \sigma^2_{A_T} / \left[\sigma^2_{A_T} + (\sigma^2_{E^*}/n)\right],$$

as $\sigma^2_{E^*} = n\left[\sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$.

In conclusion, without standardization, all four experimental set-ups would result in a different T^2, even though the underlying direct and indirect (co)variances are identical. To resolve the issue, T^2 can be standardized, allowing a fair comparison between experimental set-ups.

Box A: Individual data & unrelated birds

-Animal model

The following model is used to analyze the data:

$$P_i = Mu + FE + Animal_i + \sum_{i \neq j}^{n-1} Animal_j + Cage + Error,$$

where i refers to the animal on which the phenotype was recorded and j refers to the cage-mates. The output of the model is:

- The direct genetic variance: $\sigma^2_{A_D}$
- The direct-indirect genetic covariance $\sigma_{A_{DI}}$
- The indirect genetic variance: $\sigma^2_{A_I}$
- The cage variance: $\sigma^2_{Cage} = 2\sigma_{E_{DI}} + (n-2)\sigma^2_{E_I}$
- The error variance: $\sigma^2_E = \sigma^2_{E_D} - 2\sigma_{E_{DI}} + \sigma^2_{E_I}$

-Sire model

The following model is used to analyze the data:

$$P_i = Mu + FE + Sire_i + \sum_{i \neq j}^{n-1} Sire_j + Cage + Error.$$

The output of the model is:

- The direct genetic variance: $\sigma^2_{S_D} = \frac{1}{4}\sigma^2_{A_D}$
- The direct-indirect genetic covariance $\sigma_{S_{DI}} = \frac{1}{4}\sigma_{A_{DI}}$
- The indirect genetic variance: $\sigma^2_{S_I} = \frac{1}{4}\sigma^2_{A_I}$
- The cage variance: $\sigma^2_{Cage} = 2\sigma_{E_{DI}} + (n-2)\sigma^2_{E_I} + \frac{3}{4}\left(2\sigma_{A_{DI}} + (n-2)\sigma^2_{A_I}\right)$
- The error variance: $\sigma^2_E = \sigma^2_{E_D} - 2\sigma_{E_{DI}} + \sigma^2_{E_I} + \frac{3}{4}(\sigma^2_{A_D} - 2\sigma_{A_{DI}} + \sigma^2_{A_I})$

Phenotypic variance

For the animal model, $\sigma^2_P = \sigma^2_{A_D} + (n-1)\sigma^2_{A_I} + \sigma^2_{Cage} + \sigma^2_E$. For the sire model, $\sigma^2_P = \sigma^2_{S_D} + (n-1)\sigma^2_{S_I} + \sigma^2_{Cage} + \sigma^2_E$. Consequently, $\sigma^2_P = \sigma^2_{A_D} + (n-1)\sigma^2_{A_I} + \sigma^2_{E_D} + (n-1)\sigma^2_{E_I}$, irrespective of the model.

Box B: Individual data & half-sibs

-Animal model

The following model is used to analyze the data:

$$P_i = Mu + FE + Animal_i + \sum_{i \neq j}^{n-1} Animal_j + Cage + Error,$$

where i refers to the animal on which the phenotype was recorded and j refers to the cage-mates. The output of the model is:

- The direct genetic variance: $\sigma^2_{A_D}$
- The direct-indirect genetic covariance $\sigma_{A_{DI}}$
- The indirect genetic variance: $\sigma^2_{A_I}$
- The cage variance: $\sigma^2_{Cage} = 2\sigma_{E_{DI}} + (n-2)\sigma^2_{E_I}$
- The error variance: $\sigma^2_E = \sigma^2_{E_D} - 2\sigma_{E_{DI}} + \sigma^2_{E_I}$

-Sire model

The following model cannot be used to analyze the data:

$$P_i = Mu + FE + Sire_i + \sum_{i \neq j}^{n-1} Sire_j + Cage + Error,$$

as $Sire_i$ and $Sire_j$ refer to the same animal (cage members have the same sire) and complete confounding occurs between direct and indirect genetic effects. Alternatively, the following model can be used (as suggested in Chapter 3):

$P_i = Mu + FE + Sire_i + Cage + Error.$

The output of the model is:

- The genetic variance: $\sigma^2_{S_T} = \frac{1}{4}\sigma^2_{A_T}$
- The cage variance: $\sigma^2_{Cage} = 2\sigma_{E_{DI}} + (n-2)\sigma^2_{E_I} + \frac{3}{4}\left(2\sigma_{A_{DI}} + (n-2)\sigma^2_{A_I}\right)$
- The error variance: $\sigma^2_E = \sigma^2_{E_D} - 2\sigma_{E_{DI}} + \sigma^2_{E_I} + \frac{3}{4}(\sigma^2_{A_D} - 2\sigma_{A_{DI}} + \sigma^2_{A_I})$

Phenotypic variance

For the animal model, $\sigma^2_P = \sigma^2_{A_D} + (n-1)\sigma^2_{A_I} + \frac{1}{4}(n-1)(2\sigma_{A_{DI}} + (n-2)\sigma^2_{A_I}) + \sigma^2_{Cage} + \sigma^2_E$. For the sire model, $\sigma^2_P = \sigma^2_{S_T} + \sigma^2_{Cage} + \sigma^2_E$. Consequently, $\sigma^2_P = \frac{3}{4}\left[\sigma^2_{A_D} + (n-1)\sigma^2_{A_I}\right] + \frac{1}{4}\sigma^2_{A_T} + \sigma^2_{E_D} + (n-1)\sigma^2_{E_I}$, irrespective of the model. This is in accordance with Bergsma *et al.* (2008), who showed that for individual data on related animals $\sigma^2_P = \sigma^2_{A_D} + \sigma^2_{E_D} + (n-1)(\sigma^2_{A_I} + \sigma^2_{E_I}) + (n-1)\, r\, [2\sigma_{A_{DI}} + (n-2)\sigma^2_{A_I}]$, where r is the relatedness within a group.

Box C: Pooled data & unrelated birds

-Animal model

The following model is used to analyze the data:

$$P_k^* = Mu + FE + \sum_k^n Animal_k + Error,$$

where k refers to the cage-members. The output of the model is:

- The genetic variance: $\sigma^2_{A_T}$
- The error variance: $\sigma^2_{E^*} = n\left[\sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$

-Sire model

The following model is used to analyze the data:

$$P_k^* = Mu + FE + \sum_k^n Sire_k + Error.$$

The output of the model is:

- The genetic variance: $\sigma^2_{S_T} = \frac{1}{4}\sigma^2_{A_T}$
- The error variance: $\sigma^2_{E^*} = n\left[\sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I} + \frac{3}{4}\sigma^2_{A_T}\right]$

Phenotypic variance

For the animal model, $\sigma^2_{P^*} = n\sigma^2_{A_T} + \sigma^2_{E^*}$. For the sire model, $\sigma^2_{P^*} = n\sigma^2_{S_T} + \sigma^2_{E^*}$. Consequently, $\sigma^2_{P^*} = n\left[\sigma^2_{A_T} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$, irrespective of the model. This is in accordance with Chapter 4, where we showed that for pooled data on unrelated animals $\sigma^2_{P^*} = n\left[\sigma^2_{A_D} + 2(n-1)\sigma_{A_{DI}} + (n-1)^2\sigma^2_{A_I} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$.

Box D: Pooled data & half-sibs

-Animal model

The following model is used to analyze the data:

$$P_k^* = Mu + FE + \sum_k^n Animal_k + Error,$$

where k refers to the cage-members. The output of the model is:

- The genetic variance: $\sigma^2_{A_T}$
- The error variance: $\sigma^2_{E^*} = n\left[\sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$

-Sire model

The following model is used to analyze the data:

$P_k^* = Mu + FE + n * Sire_k + Error.$

The output of the model is:

- The genetic variance: $\sigma^2_{S_T} = \frac{1}{4}\sigma^2_{A_T}$
- The error variance: $\sigma^2_{E^*} = n\left[\sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I} + \frac{3}{4}\sigma^2_{A_T}\right]$

Alternatively, one could fit the following model to analyze the data:

$P_k^* = Mu + FE + Sire_k + Error.$

The output of the model is:

- The genetic variance: $n^2\sigma^2_{S_T} = n^2\frac{1}{4}\sigma^2_{A_T}$
- The error variance: $\sigma^2_{E^*} = n\left[\sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I} + \frac{3}{4}\sigma^2_{A_T}\right]$

Phenotypic variance

For the animal model, $\sigma^2_{P^*} = n\sigma^2_{A_T} + \left[\frac{1}{4}n(n-1)\sigma^2_{A_T}\right] + \sigma^2_{E^*}$. For the sire model, $\sigma^2_{P^*} = n^2\sigma^2_{S_T} + \sigma^2_{E^*}$. Consequently, $\sigma^2_{P^*} = n\left[\frac{1}{4}(n+3)\sigma^2_{A_T} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$, irrespective of the model.

h^2 as a measure of inheritance

The dependency of the phenotypic variance on the experimental set-up also holds true for non-social interaction traits:

- Experimental set-up A-B, individual data

As shown in Equation 1.1:

$$P_i = A_{D_i} + E_{D_i}.$$

Consequently,

$$\sigma^2_P = \sigma^2_{A_D} + \sigma^2_{E_D}.$$

- Experimental set-up C, pooled data & unrelated birds

A pooled phenotype (P^*) is obtained by summing the individual phenotypes within a group:

$$P^* = \sum_{k=1}^{n} P_k,$$

where k refers to the group members.

Based on the above formula and Equation 1.1:

$$P^* = \sum_{k=1}^{n} \left[A_{D_k} + E_{D_k}\right].$$

Consequently,

$$\sigma^2_{P^*} = n\left[\sigma^2_{A_D} + \sigma^2_{E_D}\right].$$

- Experimental set-up D, pooled data & half-sibs

As shown in the previous section:

$$P^* = \sum_{k=1}^{n} \left[A_{D_k} + E_{D_k}\right].$$

The genetic component can be written in terms of the sire, dam and Mendelian sampling term. Because all animals within a cage have the same sire, $\sum_{k=1}^{n} A_{Sire_{D_k}}$ can be replaced by $nA_{Sire_{D_k}}$:

$$P^* = \frac{1}{2} nA_{Sire_{D_k}} + \sum_{k=1}^{n} \left[\frac{1}{2} A_{Dam_{D_k}} + MS_{D_k} + E_{D_k}\right].$$

Consequently,

$$\sigma^2_{P^*} = n\left[\frac{1}{4}(n+3)\sigma^2_{A_D} + \sigma^2_{E_D}\right].$$

Experimental set-ups C and D result in a pooled phenotypic variance rather than an individual phenotypic variance. Therefore, when calculating h^2, the pooled phenotypic variance was divided by the number of group members:

- Experimental set-up A-B-C

$$h^2 = \sigma^2_{A_D} / \left[\sigma^2_{A_D} + \sigma^2_{E_D}\right]$$

- Experimental set-up D

$$h^2 = \sigma^2_{A_D} / \left[\frac{1}{4}(n+3)\sigma^2_{A_D} + \sigma^2_{E_D}\right]$$

For non-social interactions traits, experimental set-ups A, B and C would result in the same h^2. However, the pooled phenotypic variance of experimental set-up D would be larger than the pooled phenotypic variance of experimental set-up C and would be more than n times larger than the individual phenotypic variance of experimental set-up A and B. Consequently, h^2 in experimental set-up D would be lower. When similar individual phenotypes are present in the same group, the group measurements will show increased variation. This is the case for experimental set-up D, as related animals are placed in the same group.

To allow a fair comparison between the h^2 obtained in different experimental set-ups, we would have to use a formula for which the outcome is independent of the experimental set-up:

$$h^2_{standardized} = \sigma^2_{A_D} / \left[\sigma^2_{A_D} + \sigma^2_{E_D}\right].$$

In conclusion, in case the experimental set-up consists of pooled data on related birds (experimental set-up D), h^2 can differ from other experimental set-ups. To resolve the issue, h^2 should be calculated as $\sigma^2_{A_D} / \left[\sigma^2_{A_D} + \sigma^2_{E_D}\right]$, irrespective of the experimental set-up.

6.2.2 Simulation

The simulation served two purposes. First, to confirm the theoretical derivations regarding T^2 and the estimated variance components in Box A, B, C and D. Second, to give an indication of the impact of the experimental set-up on T^2. Using R v2.12.2 (Venables *et al.*, 2011), a base population of 500 sires and 6 000 dams was simulated. Each animal in the base population was assigned direct and indirect breeding values. Breeding values were drawn from a multivariate normal distribution, where the direct and indirect genetic variances were set to 1.00 and the covariance was set to 0.50. Each sire was mated to 12 randomly chosen dams resulting in 1 offspring per mating and 6 000 offspring in total. Offspring breeding values were obtained as $A = \frac{1}{2}A_{Sire} + \frac{1}{2}A_{Dam} + MS$. The direct and indirect Mendelian sampling terms were drawn from a multivariate normal distribution,

where the direct and indirect genetic variances were set to 0.50 and the covariance was set to 0.25. Each offspring was also assigned direct and indirect environmental values. The environmental values were drawn from a multivariate normal distribution where the direct and indirect environmental variances were set to 3.00 and the covariance was set to 1.50. Depending on the experimental set-up, either four unrelated animals or four half-sibs (sire related) were assigned to a group, resulting in 1 500 groups (Experimental set-up A and C *vs* B and D). Depending on the experimental set-up, either individual or pooled records were obtained (Experimental set-up A and B *vs* C and D). Individual phenotypes were obtained by using Equation 1.2. Pooled phenotypes were obtained by adding up the individual phenotypes within a group. One hundred replicates were produced. Either an animal or a sire model was used to estimate variance components. Experimental set-up A and D were analysed with an animal model. More details on the model can be found in Box A and D. Experimental set-up B and C were analysed with a sire model. More details on the model can be found in Box B and C.

The true indirect genetic and environmental effects were large. Moreover, the direct-indirect genetic and environmental correlations were positive and moderate. Consequently, experimental set-ups were expected to show clear differences. Table 6.1 shows the true and estimated variance components for each experimental set-up. In all four cases, the true values were within the 95% confidence interval of the estimates, thereby confirming the derivations made in Box A, B, C and D.

Table 6.2 shows the true and estimated T^2 (non-standardized). Large differences in T^2 were observed between experimental set-ups, thereby confirming what was expected. A higher relatedness within a group caused T^2 to decrease by 12 to 16%. A higher level of data collection (pooled data instead of individual data) caused T^2 to decrease by 70%. Based on experimental set-up A or B, the trait appears highly heritable (T^2 = 0.71-0.81). However, based on experimental set-up C or D, the trait appears moderately heritable (T^2 = 0.21-0.25).

Table 6.1 True and estimated variance components for different experimental set-ups and models

Variance components	True	Estimated
Experimental set-up A; individual data & unrelated birds & animal model		
$\sigma^2_{A_D}$	1.00	1.04 ± 0.42
$\sigma_{A_{DI}}$	0.50	0.54 ± 0.33
$\sigma^2_{A_I}$	1.00	1.03 ± 0.31
σ^2_{Cage}	9.00	8.89 ±1.25
σ^2_E	3.00	3.00 ± 0.19
σ^2_P	16.00	16.03 ± 0.48
Experimental set-up B; individual data & half-sibs & sire model		
$\sigma^2_{S_T}$	3.25	3.29 ± 0.44
σ^2_{Cage}	11.25	11.30 ± 0.49
σ^2_E	3.75	3.75 ± 0.07
σ^2_P	18.25	18.33 ± 0.56
Experimental set-up C; pooled data & unrelated birds & sire model		
$\sigma^2_{S_T}$	3.25	3.25 ± 1.29
$\sigma^2_{E^*}$	195	195.91 ± 7.34
$\sigma^2_{P^*}/4$	52	52.23 ± 1.73
Experimental set-up D; pooled data & half-sibs & animal model		
$\sigma^2_{A_T}$	13	13.07 ± 1.91
$\sigma^2_{E^*}$	156	155.42 ± 12.81
$\sigma^2_{P^*}/4$	61.75	61.74 ± 2.31

Table 6.2 True and estimated T^2 for different experimental set-ups

Experimental set-up	True T^2	Estimated T^2
A; individual data & unrelated birds	0.81	0.85 ± 0.30
B; individual data & half-sibs	0.71	0.72 ± 0.10
C; pooled data & unrelated birds	0.25	0.25 ± 0.10
D; pooled data & half-sibs	0.21	0.21 ± 0.03

6.2.3 Empirical results, this thesis

In this section, I will show the impact of the experimental set-up on T^2 for survival time. In Chapter 2 we estimated direct and indirect genetic effects for survival time in two White Leghorn pure lines (W1 and WB) and their reciprocal cross (W1xWB and WBxW1). Variance components were estimated using individual survival time data obtained from cages with unrelated birds (Experimental set-up A). Indirect genetic effects were large in purebreds and crossbreds. Direct–indirect genetic correlations were low in purebreds (0.20 ± 0.21 and -0.28 ± 0.18), but moderately to highly negative in crossbreds (-0.37 ± 0.17 and -0.83 ± 0.10). Based on the variance components obtained from individual data from cages with unrelated birds (Experimental set-up A), one can predict the variance components that would be obtained with experimental set-up B, C and D. Therefore, I could use the variance components obtained in Chapter 2 (Experimental set-up A) to calculate T^2 for all four experimental set-ups (Table 6.3). This allowed me to illustrate the dependency of T^2 for survival time on the experimental set-up.

Small differences in T^2 were observed between experimental set-ups that differed in relatedness. T^2 stayed constant or showed a small decrease with higher relatedness within a group. Large differences in T^2 were observed between experimental set-ups that differed in the level of data collection (individual data *vs* pooled data). Recording phenotypes on a group level instead of on an individual level caused T^2 to decrease by 25 up to 40%.

Table 6.3 Variance components and T^2 for survival time across four experimental set-ups for two purebred lines (W1 and WB) and their reciprocal cross (W1xWB and WBxW1)

	W1	WB	W1xWB	WBxW1
$\sigma^2_{A_D}$	656	1 400	536	997
$\sigma_{A_{DI}}$	51	-161	-197	-726
$\sigma^2_{A_I}$	100	228	536	767
σ^2_{Cage}	803	1 200	1 984	2 379
σ^2_E	7 976	12 686	11 732	15 655
σ^2_P Experimental set-up A	9 735	15 970	15 860	21 332
σ^2_P Experimental set-up B	9 962	16 0751	16 369	21 394
$\sigma^2_{P^*}/4$ Experimental set-up C	11 654	18 108	20 713	26 057
$\sigma^2_{P^*}/4$ Experimental set-up D	14 447	21 837	26 980	31 373
T^2 Experimental set-up A	0.19	0.16	0.26	0.17
T^2 Experimental set-up B	0.19	0.15	0.25	0.17
T^2 Experimental set-up C	0.14	0.12	0.18	0.12
T^2 Experimental set-up D	0.13	0.11	0.15	0.11

6.2.4 Empirical results, literature

In this section, the study by Biscarini *et al.* (2008) is used to illustrate that measures of inheritance obtained from different experimental set-ups are not directly comparable. Moreover, I will address what happens when a classic direct genetic model is used to estimate variance components for a social interaction trait (Box E and F). Box E shows the estimated variance components when individual data on unrelated birds is analyzed with a direct genetic model even though the trait is a social interaction trait. Box F shows the estimated variance components when individual data on half-sibs is analysed with a direct genetic model even though the trait is a social interaction trait. The aim is to gain insight into the underlying direct and indirect genetic (co)variances of the estimated variance components.

Box E: Individual data & unrelated birds & direct genetic model

The following model is used to analyze the data:

$P_i = Mu + FE + Animal_i + Cage + Error,$

where i refers to the animal on which the phenotype was recorded. The output of the model is:

- The direct genetic variance: $\sigma^2_{A_D}$
- The cage variance: $\sigma^2_{Cage} = 2\sigma_{E_{DI}} + (n-2)\sigma^2_{E_I} + 2\sigma_{A_{DI}} + (n-2)\sigma^2_{A_I}$
- The error variance: $\sigma^2_E = \sigma^2_{E_D} - 2\sigma_{E_{DI}} + \sigma^2_{E_I} - 2\sigma_{A_{DI}} + \sigma^2_{A_I}$

Phenotypic variance

$\sigma^2_P = \sigma^2_{A_D} + \sigma^2_{Cage} + \sigma^2_E$. Consequently, $\sigma^2_P = \sigma^2_{A_D} + (n-1)\sigma^2_{A_I} + \sigma^2_{E_D} + (n-1)\sigma^2_{E_I}$. This is the same phenotypic variance as found in Box A.

The measure of inheritance is calculated by dividing the genetic variance by the phenotypic variance. In this case, this will result in h^2.

Box F: Individual data & half-sibs & direct genetic model

The following model is used to analyze the data:

$P_i = Mu + FE + Animal_i + Cage + Error,$

where i refers to the animal on which the phenotype was recorded. The output of the model is:

- The genetic variance: $\sigma^2_A = \frac{1}{4}\sigma^2_{A_T} + \frac{3}{4}\sigma^2_{A_D}$
- The cage variance: $\sigma^2_{Cage} = 2\sigma_{E_{DI}} + (n-2)\sigma^2_{E_I} + \frac{3}{4}\left[2\sigma_{A_{DI}} + (n-2)\sigma^2_{A_I}\right]$
- The error variance: $\sigma^2_E = \sigma^2_{E_D} - 2\sigma_{E_{DI}} + \sigma^2_{E_I} + \frac{3}{4}\left[-2\sigma_{A_{DI}} + \sigma^2_{A_I}\right]$

Phenotypic variance

$\sigma^2_P = \sigma^2_A + \sigma^2_{Cage} + \sigma^2_E$. Consequently, $\sigma^2_P = \frac{3}{4}\left[\sigma^2_{A_D} + (n-1)\sigma^2_{A_I}\right] + \frac{1}{4}\sigma^2_{A_T} + \sigma^2_{E_D} + (n-1)\sigma^2_{E_I}$. This is the same phenotypic variance as found in Box B.

The measure of inheritance is calculated by dividing the genetic variance by the phenotypic variance. In this case, this will not result in h^2, nor will it result in the T^2 of experimental set-up B.

Biscarini *et al.* (2008) estimated variance components for body weight in laying hens using individual and pooled data. Individual body weight was recorded on hens housed in cages of four. Half of the cages consisted of unrelated birds, while the other half consisted of full-sibs. Individual body weight was analysed with a direct animal model. Pooled body weight was calculated by adding up the individual records within a cage. Pooled body weight was analysed with a pooled animal model. Biscarini *et al.* (2008) expected individual and pooled data to yield the same h^2. For individual data, the h^2 was calculated as $\sigma_A^2/(\sigma_A^2 + \sigma_E^2)$. For pooled data, the h^2 was calculated as $\sigma_A^2/\left[\sigma_A^2 + \left(\sigma_{E^*}^2/4\right)\right]$. Biscarini *et al.* (2008), therefore, did not take into account that the pooled phenotypic variance increased because half of the cages consisted of full-sibs.

For body weight at 43 and 51 weeks, the h^2 based on individual data was 43-53% larger than the h^2 based on pooled data. A decrease in h^2 when shifting from individual to pooled data is caused by a decrease in genetic variance and/or an increase in phenotypic variance:

- The decrease in h^2 could not be attributed to a decrease in genetic variance when shifting from individual to pooled data. On the contrary, the genetic variance for individual body weight was 24-27% larger than that for pooled body weight. There are two possible reasons for this difference. First, no cage effect was fitted in the direct animal model for individual body weight. If a random cage effect was present, it was partly absorbed by the genetic variance. Second, the genetic variance for individual body weight contains $\sigma_{A_D}^2$ and part of $\sigma_{A_T}^2$ (Box E and F), while the genetic variance for pooled body weight contains $\sigma_{A_T}^2$ (Box C and D). In most studies $\sigma_{A_T}^2$ is larger than $\sigma_{A_D}^2$, but the opposite has also been observed (Wilson *et al.*, 2011; Sartori and Mantovani,2013; Costa e Silva *et al.*, 2013)

-The decrease in h^2 could be attributed to an increase in pooled error variance that exceeded expectations. Expectations were that $\sigma_{E^*}^2$ would be four times larger than σ_E^2. However, $\sigma_{E^*}^2$ for pooled body weight was five times larger than σ_E^2 for individual body weight. In Chapter 4 we have shown that $\sigma_{E^*}^2$ for a pooled trait can be predicted based on the σ_{Cage}^2 and σ_E^2 for an individual trait, *i.e.* $\sigma_{E^*}^2 = n^2\sigma_{Cage}^2 + n\sigma_E^2$. This shows that as soon as a random cage effect is present, the pooled error variance should be more than four times larger than the individual error variance.

In conclusion, Biscarini *et al.* (2008) expected the h^2 based on individual and pooled data to be the same. However, the h^2 based on individual data was 43-53% larger than the h^2 based on pooled data. With the presence of a random cage effect and social interactions, I do not expect the h^2 based on individual and pooled data to be the same. I propose to analyse individual body weight at 43 and 51 weeks with a

direct-indirect genetic model (Box A-B, model includes a random cage effect). I expect that the estimated variance components from individual and pooled data will then no longer conflict.

6.2.5 Conclusion

The level of data collection (individual *vs* pooled data) and the within-group relatedness affects the phenotypic variance. Consequently it also affects T^2, which is the total heritable variance relative to phenotypic variance. For survival time in laying hens, a 30-40% decrease in T^2 is observed when using pooled data instead of individual data. This illustrates that T^2, as a measure of inheritance, should be used with care. When comparing T^2's obtained from different experimental set-ups, the problem can be circumvented by using a standardized T^2:
$T^2_{standardized} = \sigma^2_{A_T} / \left[\sigma^2_{A_T} + \sigma^2_{E_D} + 2(n-1)\sigma_{E_{DI}} + (n-1)^2\sigma^2_{E_I}\right]$.

6.3 Practical implementation

In this section I will discuss the practical implementation and coinciding hurdles of genetically improving social interaction traits. How can we improve survival in non-beak-trimmed group-housed laying hens? What is the impact of improper modelling? How do we improve social interaction traits when animals are kept in large groups?

6.3.1 Improving survival

Trait definition

The trait we aim to improve is survival time in non-beak-trimmed group-housed laying hens. This trait differs from survival time in beak-trimmed laying hens or survival time in individually housed laying hens. When hens are beak-trimmed or housed individually, survival time mainly reflects the robustness against bacterial, viral and parasitic diseases (Fossum *et al.*, 2009). Most pure line animals are housed individually. In addition, most crossbred group-housed animals are beak-trimmed. Because the majority of the survival time records do not reflect the consequences of feather pecking or cannibalistic behaviour, selection for improved survival time would mainly improve robustness against diseases. If we aim to decrease mortality due to feather pecking and cannibalistic behaviour, survival time in non-beak-trimmed group-housed laying hens should be treated as a separate trait.

Using feather condition score as an indicator trait

In Chapter 3 we have shown that if a line or cross has a low mortality rate, this implies low phenotypic variation for survival time and a high percentage of censored records. If a large part of the laying hens has the same phenotype, it becomes difficult to determine which bird has the best breeding value. Alternatively, feather condition score can be used (Biscarinni *et al.*, 2010; Brinker *et al.*, 2014). Nipped feathers and denuded areas on the back, rump and belly are good indicators that the bird was targeted by feather peckers and cannibals (Figure 6.2; Bilčík and Keeling, 2000). The back, rump and belly are chosen as informative areas because these areas are less affected by abrasion. Even if a line or cross has a low mortality rate, birds will still show variation in feather condition score, allowing us to determine which bird has the best breeding value.

To quantify how effective selection for improved feather condition score would be to improve survival time, I propose to estimate the (direct, indirect and total) genetic correlations between both traits.

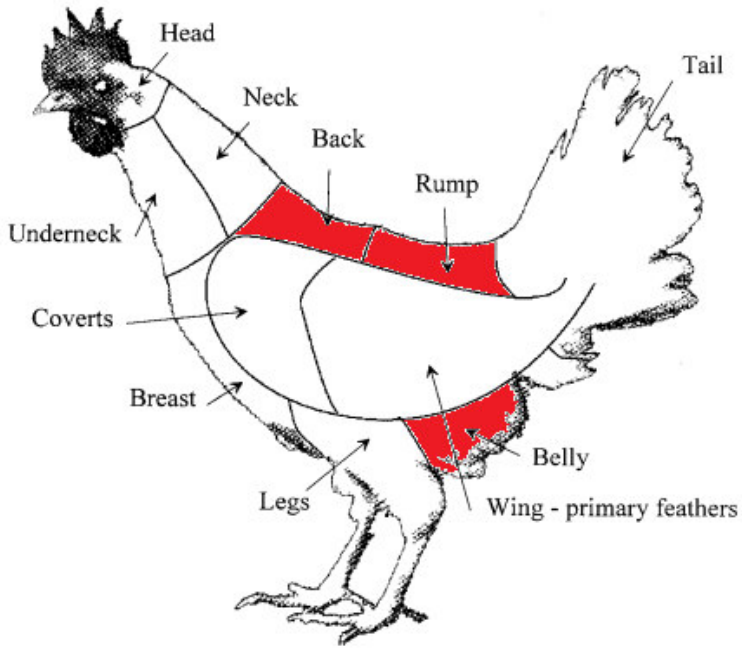

Figure 6.2 Areas where feather condition score can be recorded (Bilčík and Keeling, 2000); feather condition score in the red areas are informative indicator traits for feather pecking and cannibalistic behaviour

6.3.2 Improper modelling of social interaction traits

With social interactions, an individual has a direct (genetic) effect on its own phenotype and an indirect (genetic) effect on the phenotypes of its group mates. Traditionally, geneticists only model direct genetic effects, while the indirect genetic effects remain ignored. When group members are unrelated, a classic direct animal model would estimate an animals' direct breeding value. In Chapter 2 we have shown the consequences of selecting unrelated group members based on their estimated direct breeding value for survival time. In the pure lines, this would result in a realized heritability of 0.08 for W1 and 0.06 for WB. In the reciprocal cross, this would result in a realized heritability of 0.00 in W1xWB and -0.06 in WBxW1. Thus, crossbreds would fail to respond to selection or would respond in the opposite direction due to the strong negative direct-indirect genetic correlation. Individuals with a positive (direct) breeding value for their own survival have, on average, a negative (indirect) breeding value for the survival of their cage mates. Therefore, selecting animals with a high estimated direct breeding value would imply selecting animals that, on average, show more feather pecking and cannibalistic behaviour. To ensure positive and effective response to selection, animals should be selected based on their estimated total breeding value.

6.3.3 Increasing group size

In this thesis, we have looked at survival time in non-beak-trimmed group-housed laying hens that were housed in cages of four (Chapter 2, 3, 4) and five (Chapter 5). With the shift from battery cages to floor housing and aviaries, larger groups become reality. Often the question is raised whether the estimated direct and indirect breeding values for survival time in smaller groups can be used to approximate the estimated total breeding value for survival time in larger groups. Equation 1.4, $A_{T_i} = A_{D_i} + (n-1)A_{I_i}$, and Equation 1.5, $\sigma^2_{A_T} = \sigma^2_{A_D} + 2(n-1)\sigma_{A_{DI}} + (n-1)^2\sigma^2_{A_I}$, may suggest that the total genetic effect and total genetic variance can become very large as group size increases. However, results from smaller groups often cannot be extrapolated to larger groups (using Equation 1.4 and 1.5) due the presence of dilution and genotype by environment interaction (GxE). Dilution is the phenomenon where indirect genetic effects get smaller as groups get bigger (more details can be found in the next section). GxE might occur because small and large groups not only differ in size, but are often also in distinctly different housing systems, *e.g.* battery cages *vs* floor housing and aviaries in laying hens.

I do not expect selection for increased survival time in battery cages to yield negative response to selection in larger housing systems. It will, however, yield suboptimal response to selection, because the genetic correlation between total breeding values in small *vs* larger groups is probably smaller than one. Therefore I propose to investigate the impact of group size on indirect genetic effects further. As a first step I would advise to estimate the genetic correlations between direct, indirect and total genetic effects in cages of four and cages of 12. This will help gain insight into the magnitude and nature of indirect genetic effects in different group sizes. Accurate estimation of these correlations will require fairly large data sets. Based on the obtained knowledge of indirect genetic effects in different group sizes a selection experiment can be set up, thereby investigating whether results from smaller groups can be extrapolated to larger groups.

Dilution

As groups get bigger, an animal can probably spend less time interacting with each of its group mates. Therefore, the indirect genetic effect that an animal has on the phenotype of each of its group mates might become smaller in larger groups. The phenomenon, where indirect breeding values get smaller as groups get bigger, is known as dilution. Dilution is expected to be of importance for social interaction traits that involve physical contact between animals, *e.g.* feather pecking in laying hens, tail biting in pigs, cannibalism in fish or aggression in mink. Dilution is expected to be negligible for social interaction traits that do not involve physical contact among animals, *e.g.* alarm calling, airborne infectious diseases or traits affected by pheromones. The dilution of $\mathrm{A_I}$ can be modelled by expressing indirect genetic effects as a function of group size in the direct-indirect animal model (Bijma, 2010; Canario *et al.*, 2010):

$$\mathrm{A_I}(\mathrm{n}) = \left(\frac{\bar{\mathrm{n}} - 1}{\mathrm{n} - 1}\right)^{\mathrm{d}} \mathrm{A}_{\mathrm{I}_{\bar{\mathrm{n}}}},$$

where d is the dilution coefficient (ranging from 0 till 1), n is the group size, $\bar{\mathrm{n}}$ is the average group size and $\mathrm{A}_{\mathrm{I}_{\bar{\mathrm{n}}}}$ is the indirect genetic effect expressed in a group of averge size. When d is 0, the indirect genetic effect is independent of group size, and the total genetic effect increases with increasing group size (Equation 1.4). When d is 1, the indirect genetic effect decreases with increasing group size, and the total genetic effect remains constant.

Few studies estimated the dilution coefficient. Canario *et al.* (2010) estimated direct and indirect genetic effects for average daily gain in pigs. Group size varied between five and 15 boars. Complete dilution (d=1) gave the highest likelihood. The indirect genetic effect decreased and the total breeding value remained constant when group size increased. Duijvesteijn *et al.* (2012) estimated direct and indirect genetic effects for androstenone in boars. Androstenone is a pheromone that boars excrete to attract sows and is one of the components that causes boar taint in meat. Group size varied between three and 11 boars. Absence of dilution (d =0) gave the highest likelihood. The indirect genetic effect remained constant and the total breeding value increased as group size increased. The difference in dilution coefficient between both studies is in line with expectations when considering the biological background of the traits. For average daily gain, animals influence one another's phenotype through physical contact. As groups get bigger, animals spend less time interacting with each individual group mate. The effect that an animal had on a group members' daily gain therefore decreases with increasing group size. For androstenone, animals influence one another's phenotype without physical contact. Therefore, the effect that animals had on a group members' androstenone level was independent of group size.

For survival time in non-beak-trimmed group-housed laying hens, I expect that indirect genetic effects get smaller as groups get bigger. However, I do not expect complete dilution, as feather pecking and cannibalistic behaviour is a bigger problem in alternative housing systems, such as floor housing and aviaries, than in conventional battery-cages. In bigger groups feather peckers and cannibals have more potential victims to target. Therefore, the relative impact of indirect genetic effects on the total genetic effect will probably be larger in larger groups.

6.3.4 Conclusions

Throughout this thesis it became clear that survival time in non-beak-trimmed group-housed laying hens has substantial heritable variance. Selection for increased survival time due to reduced feather pecking and cannibalistic behaviour is possible, given a clear trait definition and the use of a proper model. With the shift from battery cages to floor housing and aviaries, more research is needed on the impact of increased group size on indirect genetic effects for survival time.

Summary

Summary

Social interactions among individuals are widespread, both in wild and domestic populations. By interacting, individuals might positively or negatively affect one another's phenotype. The effect of an animal on its own phenotype is referred to as a direct effect, while the effect of an animal on the phenotype of a conspecific is referred to as an indirect effect. These effects can have a genetic component. To genetically improve a social interaction trait, animals should be selected based on their total breeding value, which is a combination of an animal's direct and indirect breeding value. The variance of the total breeding values is known as the total genetic variance and shows the potential of a social interaction trait to respond to selection. Traditionally, geneticists often only modelled direct genetic effects, while the indirect genetic effects remained ignored. The available total heritable variation is, therefore, not exploited to its fullest.

Mortality due to feather pecking and cannibalistic behaviour is a major economic and welfare problem in non-beak-trimmed group-housed laying hens. The trait is a well-known example of a social interaction trait, as a bird's chance to survive not only depends on the tendency of the bird to become a victim, *i.e.* the direct (genetic) effect, but also on the tendency of its cage mates to be aggressors, *i.e.* the indirect (genetic) effect. The first aim of this thesis was to gain more insight into direct, indirect and total genetic effects for survival time in laying hens.

In Chapter 2, direct and indirect genetic effects were estimated for survival time in two purebred lines and their reciprocal cross. Hens were not beak-trimmed. Unrelated hens of the same line or cross were kept in groups of four. We found that indirect genetic effects contributed around half of the total heritable variation in purebreds (65% and 44%) and contributed the majority of the total heritable variation in crossbreds (87% and 72%). The direct-indirect genetic correlations were close to zero in purebreds (0.20 ± 0.21 and -0.28 ± 0.18) and moderately to highly negative in crossbreds (-0.37 ± 0.17 and -0.83 ± 0.10). Based on these estimates, it is predicted that crossbreds would fail to respond to selection or would respond in the opposite direction. We also estimated the genetic correlations between crosses. The direct genetic correlation was high (0.95 ± 0.23), whereas the indirect genetic correlation was moderate (0.42 ± 0.26). Thus, for indirect genetic effects, it mattered which parental line provided the sire and which provided the dam. This indirect parent-of-origin effect appeared to be paternally transmitted and is probably Z-chromosome linked.

In Chapter 3, total genetic effects were estimated for survival time in the crossbred offspring of three sire lines using single-step genomic BLUP. Hens were not beak-trimmed. Paternal half sibs of the same cross were kept in groups of five and no dam information was available. This data structure did not allow the estimation of direct and indirect genetic effects, but total genetic effects could be estimated. We also investigated whether the use of genotype information on the sires would increase the accuracy of their estimated breeding value. Cross validation showed that genotyped sires without progeny information had a 20 up to 110% increase in the accuracy of their estimated total breeding value, showing the added value of genomic selection.

Survival time in laying hens is recorded on individual birds. However, most production traits in laying hens are recorded at group level, resulting in pooled rather than individual records. The second aim of this thesis was to evaluate whether and how pooled data can be used to analyse social interaction traits. It was unclear whether indirect genetic effects can be estimated from pooled records (Chapter 4). Moreover, it was unclear how individual and pooled records on social interaction traits can be modelled multivariately (Chapter 5).

In Chapter 4, we investigated whether data collected at group level (pooled data) can be used to estimate direct, indirect and total genetic variances. In addition, we determined the optimal group composition, *i.e.* the optimal number of families represented in a group to minimize the standard error of the estimates. Through theoretical derivations and simulations we showed that the total genetic variance can be estimated from pooled data, but the underlying direct and indirect genetic (co)variances cannot. Moreover, we showed that with pooled data the most accurate estimate of the total genetic variance is obtained when group members belong to the same family. While quantifying the direct and indirect genetic effects is interesting from a biological perspective, estimating the total genetic effect is most important for genetic improvement. Therefore, when it is too difficult or expensive to obtain individual data, pooled data can be used to improve social interaction traits.

In Chapter 5, we aimed to find a model that yields unbiased genetic parameter estimates for a bivariate analysis of individual and pooled data on social interaction traits. Bivariate analysis of individual and pooled data can introduce an issue with the non-genetic correlation between both traits. For individual data, direct and indirect environmental effects are captured by two model terms: a random group effect and the residual. For pooled data, however, direct and indirect environmental effects are captured by the residual only. Most statistical software programs cannot fit a correlation between the random group effect of the

individual trait and the residual of the pooled trait. This can result in biased genetic parameter estimates. We showed that, to obtain unbiased genetic parameter estimates, a random group effect for the pooled trait needs to be added and correlated to the random group effect of the individual trait. The variance of the random group effect of the pooled trait needs to be set to a fixed value to avoid over-parameterization.

The general discussion (Chapter 6) addressed two topics. First, I discussed the interpretation of T^2 as a measure of inheritance for social interaction traits. For non-social interaction traits, the heritability (h^2) is used as a measure of inheritance, where the direct genetic variance is expressed relative to the phenotypic variance. Analogously, for social interaction traits, T^2 is used as a measure of inheritance, where the total genetic variance is expressed relative to the phenotypic variance. Throughout this thesis it became clear that, for social interaction traits, the level of data collection (individual *vs* pooled data) and the within-group relatedness affect the phenotypic variance. Consequently it also affects T^2. Therefore, T^2 can differ between experimental set-ups, even though the underlying genetic parameters are the same. For survival time in laying hens, a 30 up to 40% decrease in T^2 is observed when using pooled data instead of individual data. This illustrates that T^2, as a measure of inheritance, should be used with care. Second, I discussed the practical implementation and coinciding hurdles of genetically improving social interaction traits, where one of the hurdles will be improving survival time in housing systems where animals are kept in larger groups.

References

A

Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor, TJ: Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy. Sci.* 2010, 93: 743-752.

Albertsdóttir E, Eriksson S, Sigurdsson Á, Árnason T: Genetic analysis of 'breeding field test status' in Icelandic horses. *J. Anim. Breed. Genet.* 2011, 128: 124-132.

Alemu SW, Calus MPL, Muir WM, Peeters K, Vereijken A, Bijma P: Genomic prediction of survival time in a population of brown laying hens showing cannibalistic behavior. *Submitted*.

Anang A, Mielenz N, Schüler L: Genetic and phenotypic parameters for monthly egg production in White Leghorn hens. *J. Anim. Breed. Genet.* 2000, 117: 407-415.

Arnason T: Genetic evaluation of Swedish standard-bred trotters for racing performance traits and racing status. *J. Anim. Breed. Genet.* 1999, 116: 387-398.

Arnold AP, Itoh Y, Melamed E: A bird's-eye view of sex chromosome dosage compensation. *Annu. Rev. Genomics Hum. Genet.* 2008, 9: 109-127.

B

Bergsma R, Ka nis E, Knol EF, Bijma P: The contribution of social effects to heritable variation in finishing traits of domestic pigs (*Sus scrofa*). *Genetics* 2008, 178: 1559-1570.

Besbes B, Ducrocq V, Foulley JL, Protais M, Tavernier A, Tixier-Boichard M, Beaumont C: Estimation of genetic parameters of egg production traits of laying hens by restricted maximum likelihood applied to a multiple-trait reduced animal model. *Genet. Sel. Evol.* 1992, 24: 539-552.

Bessei W: Untersuchungen zur Heritabilität des Federpickverhaltens bei Junghennen. *Arch. Geflügelk.* 1984, 48: 224-231.

Bhagwat AL, Craig JV: Selecting for age at first egg: Effects on social dominance. *Poult. Sci.* 1997, 56: 361-363.

Bijma P, Muir WM, van Arendonk JAM: Multilevel selection 1: quantitative genetics of inheritance and response to selection. *Genetics* 2007, 175: 277-288.

Bijma P, Wade MJ: The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *J. Evol. Biol.* 2008, 21: 1175-1188.

Bijma P: Multilevel selection 4: modeling the relationship of indirect genetic effects and group size. *Genetics* 2010, 186, 1029-1031.

Bijma P: A general definition of the heritable variation that determines the potential of a population to respond to selection. *Genetics* 2011a, 189: 1347-1359.

Bijma P: Breeding for social interaction, for animal welfare. In *Encyclopedia of Sustainability Science and Technology*, Meyers RA (ed.), Springer Science and Business Media LLC, 2011b, 1477-1513.

Bilčík B, Keeling LJ: Relationship between feather pecking and ground pecking in laying hens and the effect of group size. *Appl. Anim. Behav. Sci.* 2000, 68: 55-66.

Biscarini F, Bovenhuis H, van Arendonk JAM: Estimation of variance components and prediction of breeding values using pooled data. *J. Anim. Sci.* 2008, 86: 2845-2852.

Biscarini F, Bovenhuis H, van der Poel J, Rodenburg TB, Jungerius AP, van Arendonk JAM: Across-line SNP association study for direct and associative effects on feather damage in laying hens. *Behav. Genet.* 2010a, 40: 715-727.

Biscarini F, Bovenhuis H, Ellen ED, Addo S, van Arendonk JAM: Estimation of heritability and breeding values for early egg production in laying hens from pooled data. *Poult. Sci.* 2010b, 89: 1842-1849.

Blokhuis HJ, Wiepkema PR: Studies of feather pecking in poultry. *Vet. Q.* 1998, 20: 6-9.

Blokhuis HJ, Beutler A: Feather pecking damage and tonic immobility response in two lines of White Leghorn hens. *J. Anim. Sci.* 1992, 70: 170.

Bouwman AC, Bergsma R, Duijvesteijn N, Bijma P: Maternal and social genetic effects on average daily gain of piglets from birth until weaning. *J. Anim. Sci.* 2010, 88: 2883-2892.

Brinker T, Bijma P, Visscher J, Rodenburg TB, Ellen ED. Plumage condition in laying hens: genetic parameters for direct and indirect effects in two purebred layer lines. *Genet. Sel. Evol.* 2014, 46: 33.

C

Canario L, Lundeheim N, Bijma P: Pig growth is affected by social genetic effects and social litter effects that depend on group size. In *Proc. 9th World Congr. Genet. Appl. Livest. Prod.,* Leipzig, Germany 2010, 87.

Chen CY, Kachman SD, Johnson RK, Newman S, Van Vleck LD: Estimation of genetic parameters for average daily gain using models with competition effects. *J. Anim. Sci.* 2008, 86: 2525-2530.

Chen CY, Johnson RK, Newman S, Kachman SD, Van Vleck LD: Effects of social interactions on empirical responses to selection for average daily gain of boars. *J. Anim. Sci.* 2009, 87:844-849.

Cheverud JM: Evolution by kin selection: A quantitative genetic model illustrated by maternal performance in mice. *Evolution* 1984, 38: 766-777.

Christensen OF, Lund MS: Genomic relationship matrix when some animals are not genotyped. In *Proc. 60th Annual Meeting EAAP*, Barcelona, Spain 2009, No 28-12.

Costa e Silva J, Potts BM, Bijma P, Kerr RJ and Pilbeam DJ: Genetic control of interactions among individuals: contrasting outcomes of indirect genetic effects arising from neighbor disease infection and competition in a forest tree. *New Phytol.* 2003, 197: 631-641.

Council Directive 1999/74/EC: Laying down minimum standards for the protection of laying hens. *Official Journal L* 2009, 203: 53-57.

Craig JV, Jan ML, Polley CR, Bhagwat AL, Dayton AD: Change in relative aggressiveness and social dominance associated with selection for early egg production in chickens. *Poult. Sci.* 1975, 54: 1647-1658.

Craig DM: Group selection versus individual selection: an experimental analysis. *Evolution* 1982, 36: 271-282.

Craig JV, Muir WM: Group selection for adaptation to multiple-hen cages: beakrelated mortality, feathering and bodyweight responses. *Poult. Sci.* 1996, 75: 294-302.

Cuthbertson GJ: Genetic variation in feather-pecking behavior. *Br. Poult. Sci.* 1980: 21: 447-450.

D

Dematawega CMB, Berger PJ: Genetic and phenotypic parameters for 305-day yield, fertility and survival in Holsteins. *J. Dairy Sci.* 1998, 81: 2700-2709.

Denison RF, Kiers ET, West SA: Darwinian agriculture: When can humans find solutions beyond the reach of natural selection? *Q. Rev. Biol.* 2003, 78: 145-168.

Dickerson GE: Composition of hog carcasses as influenced by heritable differences in rate and economy of gain. *Iowa Agric. Exp. Stn. Res. Bull.* 1947, 345: 489-524

Ducro BJ: Accounting for preselection in genetic evaluation of dressage performance of Dutch Warmblood Horses. In *Proc. 9th World Congr. Genet. Appl. Livest. Prod.*, Leipzig, Germany 2010, 58.

Duijvesteijn N, Knol EF, Bijma P: Direct and associative effects for androstenone and genetic correlations with backfat and growth in entire male pigs. *J. Anim. Sci.* 2012, 90: 2465-2475.

E

Eaglen SAE, Bijma P: Genetic parameters of direct and maternal effects for calving ease in Dutch Holstein-Friesian cattle. *J. Dairy Sci.* 2009, 92: 2229-2237.

Ellen ED, Visscher J, van Arendonk JAM, Bijma P: Survival of laying hens: genetic parameters for direct and associative effects in three purebred layer lines. *Poult. Sci.* 2008, 87: 233-239.

Ellen ED: Genetics of survival in cannibalistic laying hens: The contribution of social effects. *Dissertation, Wageningen University, The Netherlands* 2009.

Ellen ED, Ducrocq V, Ducro BJ, Veerkamp RF, Bijma P: Genetic parameters for social effects on survival in cannibalistic layers: Combining survival analysis and a linear animal model. *Genet. Sel. Evol.* 2010, 42: 27.

Ellen ED, Rodenburg TB, Visscher J, Bijma P: The consequence of selection on social genetic effects for survival in laying hens. In *Proc. 8th European Symp. Poult. Genet.* Venice, Italy 2013, 35-38.

Ellen ED, Rodenburg BT, Albers GAA, Bolhuis JE, Camerlink I, Duivesteijn N, Knol, E, Muir WM, Peeters K, Reimert I, Sell-Kubiak E, van Arendonk JAM, Visscher J, Bijma P: The prospects of selection for social genetic effects to improve animal welfare in livestock. *Front. Genet.* 2014, 5: 377.

F

Fairfull RW, Gowe RS, Emsley JAB: Diallel cross of six long-term selected leghorn strains with emphasis on heterosis and reciprocal effects. *Br. Poult. Sci.* 1983, 24: 133-158.

Fairfull RW, Gowe RS: Use of breed resources for poultry egg and meat production. In *3rd World Congr. Genet. Appl. Livest. Prod.*, Lincoln, NE, USA 1986, 242-256.

Falconer DS, Mackay TFC: Introduction to Quantitative Genetics. *4th edition, Longman, Essex,* 1996.

Fisher RA: The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edin.* 1918, 52: 399-433.

Fossum O, Jansson DS, Etterlin PE, Vågsholm I: Causes of mortality in laying hens in different housing systems in 2001 to 2004. *Acta. Vet. Scand.* 2009, 1: 3.

Frank SA: All of life is social. *Curr. Biol.* 2007, 17: 648-650.

G

Garfied AS, Cowley M, Smith FM, Moorwood K, Stewart-Cox JE, Gilroy K, Baker S, Xia J, Dalley JW, Hurst LD, Wilkinson LS, Isles AR, Ward A: Distinct physiological and behavioural functions for parental alleles of imprinted Grb10. *Nature* 2011, 469: 534-538.

Gatewood JD, Wills A, Shetty S, Xu J, Arnold AP, Burgoyne PS, Rissman EF: Sex chromosome complement and gonadal sex influence aggressive and parental behaviors in mice. *J. Neuro. Sci.* 2006, 26: 2335-2342.

Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R: ASReml user guide, Release 3.0. *VSN Int. Ltd., Hemel Hempstead* 2009.

Goodnight CJ: The influence of environmental variation on group and individual selection in a cress. *Evolution* 1985, 39: 545-558.

Griffin AS, West SA, Buckling A: Cooperation and competition in pathogenic bacteria. *Nature* 2004, 430: 1024-1027.

Griffing B: Selection in reference to biological groups. I. Individual and group selection applied to populations of unordered groups. *Aust. J. Biol. Sci.* 1967, 20: 127-139.

H

Hamilton WD: The genetical evolution of social behaviour. *Theoret. I J. Biol.* 1964a, 7: 1-16.

Hamilton WD: The genetical evolution of social behaviour. *Theoret. II J. Biol.* 1964b, 7: 17-52.

Henderson CR: Best linear unbiased estimation and prediction under a selection model. *Biometrics* 1975, 31: 423-447.

Hill WG, Goddard ME, Visscher PM: Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* 2008, 4: e1000008.

Hsu WL, Johnson RK, Van Vleck LD: Effect of pen mates on growth, backfat depth and longissimus muscle area of swine. *J. Anim. Sci.* 2010, 88: 895-902.

Huber-Eicher B, Sebö F: The prevalence of feather pecking and development in commercial flocks of laying hens. *Appl. Anim. Behav. Sci.* 2011, 54: 861-873.

Hughes BO, Duncan IJH: The influence of strain and environmental factors upon feather pecking and cannibalism in fowls. *Br. Poult. Sci.* 1972, 13: 525-547.

Hughes BO: The effect of implanted gonadal hormones on feather pecking and cannibalism in pullets. *Br. Poult. Sci.* 1973, 14: 341-348.

J

Jensen P, Keeling L, Schütz K, Andersson L, Mormède P, Brändström H, Forkman B, Kerje S, Fredriksson R, Ohlsson C, Larsson S, Mallmin H, Kindmark A: Feather pecking in chickens is genetically related to behavioural and developmental traits. *Physiol. Behav.* 2005, 86: 52-60.

Jones RB, Blokhuis HJ, Beuving G: Open-field and tonic immobility responses in domestic chicks of two genetic lines differing in their propensity to feather peck. *Br. Poult. Sci.* 1995, 36: 525-530.

K

Keeling L, Jensen P: Do feather pecking and cannibalistic hens have different personalities? *Appl. Anim. Behav. Sci.* 1995, 44: 265.

Kirkpatrick M, Lande R: The evolution of maternal characters. *Evolution* 1989, 43: 485-503.

Kjaer JB, Sorensen P: Feather pecking behavior in White leghorns, a genetic study. *Br. Poult. Sci.* 1997, 38, 333-341.

Kjaer JB, Vestergaard KS: Development of feather pecking in relation to light intensity. *Appl. Anim. Behav. Sci.* 1999, 62: 243-254.

Knol EF, Leenhouwers JI, van der Lende T: Genetic aspects of piglet survival. *Livest. Prod. Sci.* 2002, 78: 47-55.

Koerhuis ANM, Thompson R: Models to estimate maternal effects for juvenile body weight in broiler chickens. *Genet. Sel. Evol.* 1997, 29: 225-249.

Kyriakou DT, Fasoulas AC: Effects of competition and selection pressure on yield response in winter rye (*Secale cereale L.*). *Euphytica* 1985, 35: 883-895.

L

Leeson S, Morrison WB: Effect of feather cover on feed efficiency in laying birds. *Poult. Sci.* 1978, 57: 1094-1096.

Ledur MC, Liljedahl LE, McMillan I, Asselstine L, Fairfull RW: Genetic effects of aging on fitness and nonfitness traits in laying hens housed three per cage. *Poult. Sci.* 2002, 82: 1223-1234.

Legarra A, Aguilar I, Misztal I: A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 2009, 92: 4656-4663.

Lowry DC, Abplanalp H: Social dominance difference, given limited access to common food, between hens selected and unselected for increased egg production. *Br. Poult. Sci.* 1972, 13: 365-376.

Lynch M, Walsh JB: Genetics and analysis of quantitative traits. *Sinauer Associates Inc., Sunderland* 1998.

M

Meuwissen TH, Hayes BJ, Goddard ME: Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 2001, 157: 1819-1829.

Misztal I, Tsuruta S, Strabel T, Auvray B , Druet T, Lee DH: Blupf90 and related programs (BGF90). In *Proc. 7th World Congr. Genet. Appl. Livest. Prod.*, Montpellier, France 2002, No 28-07.

Misztal I, Legarra A, and Aguilar I: Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *J. Dairy Sci.* 2009, 92: 4648-4655.

Mohammed HH, Grashorn MA, Bessei W: The effects of lighting conditions on the behaviour of laying hens. *Archiv für Geflügelkunde* 2010, 74: 197-202.

Moore AJ, Brodie ED, Wolf JB: Interacting phenotypes and the evolutionary process: I. Direct and indirect genetic effects of social interactions. *Evolution* 1997, 51: 1352-1362.

Mousseau TA, Fox CW: Maternal Effects as Adaptations. *Oxford University Press, New York* 1998.

Muir WM: Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poult. Sci.* 1996, 75: 447-458.

Muir WM: Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* 2005, 170: 1247-1259.

Muir WM, Bijma P, Schinckel A: Multilevel selection with kin and non-kin groups, experimental results with Japanese quail (*Coturnix japonica*). *Evolution* 2013, 67: 1598-1606.

N

Nielsen HM, Monsen BB, Ødegård J, Bijma P, Damsgård B, Toften H, et al.: Direct and social genetic parameters for growth and fin damage traits in Atlantic cod (*Gadus morhua*). *Genet. Sel. Evol.* 2014, 46: 5.

Nicol CJ, Bestman M, Gilani A-M, De Haas EN, De Jong IC, Lambton S, Wagenaar JP, Weeks CA, Rodenburg TB: The prevention and control of feather pecking: application to commercial systems. *World. Poult. Sci. J.* 2013, 69: 775-788.

Nurgiartiningsih VM, Mielenz N, Preisinger R, Schmutz M, Schueler L: Estimation of genetic parameters based on individual and group mean records in laying hens. *Br. Poult. Sci.* 2004, 45: 604-610.

O

Olson KM, Garrick DJ, Enns RM: Predicting breeding values and accuracies from group in comparison to individual observations. *J. Anim. Sci.* 2006, 84: 88-92.

P

Peeters K, Eppink TT, Ellen ED, Visscher J, Bijma P: Indirect genetic effects for survival in domestic chicken (*Gallus gallus*) are magnified in crossbred genotypes and show a parent-of-origin effect. *Genetics* 2012, 192: 705-713.

Peeters K, Ellen ED, Bijma P: Using pooled data to estimate variance components and breeding values for traits affected by social interactions. *Genet. Sel. Evol.* 2013, 45: 27.

Q

Quinton CD, Wood BJ, Miller SP: Genetic analysis of survival and fitness in turkeys with multiple-trait animal models. *Poult. Sci.* 2011, 90: 2479-2486.

R

Reik W, Walter J: Genomic imprinting: parental influence on the genome. *Nature* 2001, 2: 21-31.

Robertson A, Hill WG, Ewens WJ: Population and quantitative genetics of many linked loci in finite populations. *Proc. R. Soc. Lond.* 1983, 219: 253-264.

Rodenburg TB, Buitenhuis AJ, Ask B, Uitdehaag KA, Koene P, van der Poel JJ, Bovenhuis H: Heritability of feather pecking and open-field response of laying hens at two different ages. Poult. Sci. 2003, 82: 861-867.

Rodenburg T B, van Hierden YM, Riedstra B, Koene P, Korte SM, van der Poel JJ, Groothuis TGG, Blokhuis H: Feather pecking in laying hens: new insights and directions for research? *Appl. Anim. Behav. Sci.* 2004, 86: 291-298.

Rodenburg TB, de Haas, EN, Nielsen BL, Buitenhuis AJ: Fearfulness and feather damage in laying hens divergently selected for high and low feather pecking. *Appl. Anim. Behav. Sci.* 2010a, 128: 91-96.

Rodenburg TB, De Reu K, Tuyttens FAM: Performance, welfare, health and hygiene of laying hens in non-cage systems in comparison with cage systems. In *Alternative systems for poultry - Health, welfare and productivity*. Sandilands V, Hocking PM (ed.), CABI Publishing, 2010b, 210-224.

Rodenburg TB, Van Krimpen MM, De Jong IC, De Haas EN, Kops MS, Riedstra BJ, Nordquist RE, Wagenaar JP, Bestman M, Nicol CJ: The prevention and control and feather pecking in laying hens: identifying the underlying principles. *World Poult. Sci. J.* 2013, 69: 361-374.

S

SAS Institute Inc.: SAS/STAT User's Guide. *Version 9.1, Cary, NC* 2003.

Sartori C, Mantovani R: Indirect genetic effects and the genetic bases of social dominance: evidence from cattle. *Heredity* 2013, 110: 3-9.

Savory CJ: Feather pecking and cannibalism. *World Poult. Sci. J.* 1995, 51: 215-219.

Schaeffer L: Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 2006, 123: 218-223.

T

Tuiskula-Haavisto, M, de Koning, DJ, Honkatukia, M, Schulman, NF, Mäki-Tanila A Quantitative trait loci with parent-of-origin effects in chicken. *Genet. Res.* 2004, 84: 57-66.

Tuiskula-Haavisto M, Vilkki J: Parent-of-origin specific QTL – a possibility towards understanding reciprocal effects in chicken and the origin of imprinting. *Cytogenet. Genome Res.* 2007, 117: 305-312.

U

Úbeda F, Gardner A: A model for genomic imprinting in the social brain: Juveniles. *Evolution* 2010, 64: 2587-2600.

V

Van Horne PLM, Achterbosch TJ: Animal welfare in poultry production systems: Impact of EU standards on world trade. *World. Poult. Sci. J.* 2008, 64: 40-52.

Van Krimpen MM, Kwakkel RP, Reuvekamp BFJ, van der Peet-Schwering CMC, den Hartog LA, Verstegen MWA: Impact of feeding management on feather pecking in laying hens. *World Poult. Sci. J.* 2005, 61: 663-686.

Van Vleck LD, Cundiff LV, Koch RM: Effect of competition on gain in feedlot bulls from Hereford selection lines. *J. Anim. Sci.* 2007, 85: 1625-1633.

VanRaden PM: Efficient methods to compute genomic predictions. *J. Dairy Sci.* 2008, 91: 4414-4423.

Venables WN, Smith DM, R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2011.

W

Wade MJ: An experimental study of group selection. *Evolution* 1977, 31: 134-153.

Wei M, van der Steen HAM, van der Werf JHJ, Brascamp EW: Relationship between purebred and crossbred parameters. I. Variances and covariances under the one-locus model. *J. Anim. Breed. Genet.* 1991, 108: 253-261.

Willham RL: The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics* 1963, 19: 18-27.

Wilson AJ, Morrissey MB, Adams MJ, Walling CA, Guinness FE, Pemberton JM, Clutton-Brock TH, Kruuk LE: Indirect genetics effects and evolutionary constraint: ananalysis of social dominance in red deer, *Cervus elaphus*. *J. Evol. Biol.* 2011, 24: 772-783.

Wolc A, Lisowski M, Szwaczkowski T: Heritability of egg production in laying hens under cumulative, multitrait and repeated measurement animal models. *Czech J. Anim. Sci.* 2007, 52: 254-259.

Wolf JB, Brodie ED, Cheverud JM, Moore AJ, Wade MJ: Evolutionary consequences of indirect genetic effects. *Trends Ecol. Evol.* 1998, 13: 64-69.

X

Xu J, Burgoyne PS, Arnold AP: Sex differences in sex chromosome gene expression in mouse brain. *Hum. Mol. Gen.* 2002, 11: 1409-1419.

Curriculum vitae

Katrijn Peeters was born on November 6th 1985 in Lokeren, Belgium. From kindergarten up to high school she went to the Berkenboom school in Sint-Niklaas. There, she obtained her high school degree in 2003. That same year she enrolled in the Agricultural and Biotechnology BSc program at the KaHo Sint-Lieven in Sint-Niklaas. She specialised in Animal Care and graduated in 2006. For her master's degree she moved to the Netherlands and enrolled in the Animal Science MSc program at Wageningen University. She specialised in Animal Breeding and Genetics (ABG) as well as Quantitative Veterinary Epidemiology (QVE). For her major ABG thesis she estimated genetic parameters for dressage performance in Belgian warmblood horses, which was in collaboration with the University of Leuven. For her major QVE thesis she identified risk factors for peripartum mastitis in Belgian dairy heifers, which was in collaboration with the University of Ghent. After graduating with a double master degree in 2009, she was accepted as a PhD candidate at the Animal Breeding and Genomics Center of Wageningen University. She worked on the STW-project entitled: "Genetics of social interactions in livestock: Improving health, welfare and productivity". Within this project she focused on improving survival and productivity in laying hens. The results of the project are presented in this thesis. In 2013, she has spent a five month period at the Department of Animal and Dairy Science of the University of Georgia (USA) as a visiting PhD student. In 2014, Katrijn started working at the Research and Technology Center of Hendrix Genetics as a Research Geneticist. She does research for ISA (layer breeding), Hybrid (turkey breeding), Hypor (pig breeding) and Troutlodge-Landcatch (aquaculture breeding). In 2014, she has spent a three month period at Landcatch (Scotland) as a visiting Research Geneticist. Currently, she is stationed at the head office in Boxmeer (The Netherlands).

Publication list

Peered reviewed papers

Peeters K, Eppink TT, Ellen ED, Visscher J, Bijma P: Indirect genetic effects for survival in domestic chickens (*Gallus gallus*) are magnified in crossbred genotypes and show a parent-of-origin effect. *Genetics* 2012, 192: 705-713.

Peeters K, Ellen ED, Bijma P: Using pooled data to estimate variance components and breeding values for traits affected by social interactions. *Genet Sel. Evol.* 2013, 45: 27.

Ellen ED, Rodenburg TB, Albers GAA, Bolhuis JE, Camerlink I, Duijvesteijn N, Knol EF, Muir WM, Peeters K, Reimert I, Sell-Kubiak E, van Arendonk JAM, Visscher J, Bijma P: The prospects of selection for social genetic effects to improve welfare and productivity in livestock. *Front. Genet.* 2014, 5: 377.

Manuscripts in preparation

Peeters K, Visscher J, Bijma P: Bivariate analysis of individual and pooled data on social interaction traits: application to survival time and early egg production in laying hens. *Submitted*.

Ellen ED, Verhoeven M, Peeters K, Gols R, Harvey JA, Wade MJ, Dicke M, Bijma P: Direct and indirect genetic effects in life history traits of flour beetles (*Tribolium castaneum*). *Submitted.*

Peeters K, Fragomeni BO, Bijma P, Muir WM, Misztal I: Single-step GBLUP for survival time in crossbred laying hens. *In preparation*.

Alemu SW, Calus MPL, Muir WM, Peeters K, Vereijken A, Bijma P: Genomic prediction of survival time in a population of brown laying hens showing cannibalistic behavior. *In preparation.*

Conference proceedings

Peeters K, Eppink TT, Ellen ED, Visscher J, Bijma P: Survival in laying hens: genetic parameters for direct and associative effects in the reciprocal crosses of two purebred layer lines. 9th WCGALP, Leipzig, Germany, 2010.

Peeters K, Eppink TT, Ellen ED, Visscher J, Bijma P: Parent-of-origin found in the associative genetic effect for survival in crossbred laying hens. 64th EAAP, Stavanger, Norway, 2011.

Peeters K, Eppink TT, Ellen ED, Visscher J, Bijma P: Direct and indirect genetic effects for survival in purebred and crossbred laying hens. 4th ICQG, Edinburgh, UK, 2012.

Peeters K, Ellen ED, Bijma P: Using pooled data to estimate variance components and breeding values for social interaction traits. 65th EAAP, Nantes, France, 2013.

Peeters K, Visscher J, Bijma P: Bivariate analysis of individual survival data and pooled early egg production data on crossbred laying hens. 10th WCGALP, Vancouver, Canada, 2014.

Peeters K, Visscher J, Bijma P: Bivariate analysis of individual and pooled data on social interaction traits. 66th EAAP, Warsaw, Poland, 2015.

Training and supervision plan

Training and supervision plan

The basic package (3 ECTS)

-WIAS introduction course	2009
-Ethics and philosophy in life sciences	2013

Scientific exposure (19 ECTS)

International conferences

-9th WCGALP, Leipzig, Germany	2010
-61th EAAP, Crete, Greece	2010
-62th EAAP, Stavanger, Norway	2011
-4th ICQG, Edinburgh, UK	2012
-ADSA-ASAS 2013 Joint Annual Meeting, Indianapolis, USA	2013
-10th WCGALP, Vancouver, Canada	2014

Seminars and workshops

-Seminar 'Friends or Fiends? Consequences of social interactions for artificial breeding programs and evolution in natural populations', Wageningen	2009
-Seminar 'Developments in genome-wide evaluation and genomic selection', Wageningen	2009
-F&G Connection Days, Vught	2010, 2012, 2014
-WIAS Science Day, Wageningen	2010-2013

Presentations

-9th WCGALP, Leipzig, Germany, oral	2010
-61th EAAP, Stavanger, Norway, poster	2011
-4th ICQG, Edinburgh, UK, oral	2012
-4th ICQG, Edinburgh, UK, poster	2012
-F&G Connection Days, Vught, oral	2012
-Hendrix Genetics Academy days, Boxmeer, oral	2012
-WIAS Science Day, Wageningen, oral	2013
-Seminar 'Genetics of social life: Agriculture meets evolutionary biology', Wageningen, oral	2013
-10th WCGALP, Vancouver, Canada, poster	2014

In-depth studies (12 ECTS)

Disciplinary and interdisciplinary courses	
-Introduction to R for statistical analysis	2009
-Genomic selection in animal breeding	2010
-Quantitative genetics, with a focus on selection theory	2010
-Evolutionary genetic approaches to study social evolution	2010
-Genomic selection in livestock	2011
-Survival analysis	2011
PhD students' discussion groups	
-Quantitative genetics discussion group (QDG), ABGC	2009-2013

Professional skills support courses (5 ECTS)

-Techniques for writing and presenting a scientific paper	2010
-Presentation skills	2012
-French course	2012
-Cross cultural communication	2014
-Project management	2015

Research skills training (2 ECTS)

-External training period, University of Georgia	2013

Didactic skills training (5 ECTS)

Supervising practicals	
-Genetic improvement of livestock	2010-2011
Supervising theses	
-MSc major	2010

Management skills training (2 ECTS)

Organisation of seminars and courses	
-Quantitative genetics, with a focus on selection theory	2010

Education and training total	**49 ECTS**

Acknowledgments

When I came to the Netherlands 9 years ago, I quickly realised that Wageningen is everything but a typical Dutch city, as it felt more like the world's melting pot. It has been such a valuable experience to live in a city where multiculturalism works, knowing that I grew up in a city where multiculturalism is the biggest issue. Those 9 years flew by and I'm happy that I've stuck by a bit longer than the initial 2-year plan...

Piter, zonder jouw vertrouwen had ik dit avontuur nooit kunnen meemaken. Bedankt om me de kans te geven om een PhD te starten. Ik heb altijd met veel plezier gewerkt aan je 'sociale interactie'-onderwerp. Jouw onderwerp, maar je liet me de vrijheid om er mijn project van te maken! Al moest je me af en toe terug fluiten, met m'n over-nieuwsgierig/perfectionistisch/koppig karakter. Aard van't beestje ;) Ik heb enorm veel van je geleerd en had me geen fijnere begeleider kunnen wensen! Bedankt!
Johan, bedankt dat ik deel mocht uitmaken van je ABGc-team. Ik ben elke dag met veel plezier naar Zodiac gekomen. Bedankt om me bij te sturen wanneer het nodig was en om me er attent op te maken wat de hoofdzaak was als ik me blind staarde op de bijzaak. Ik kijk er naar uit om straks terug samen te kunnen werken.

I would like to thank Hendrix Genetics, not only for providing the data, but also for their overall involvement in the project. A special thanks goes out to Jeroen and my RTC colleagues. Jeroen, bedankt om maanden lang elke vrijdag een momentje voor me vrij te maken en om mee te sparren over hoe we de resultaten kunnen implementeren. RTC colleagues, thank you for helping me where possible and for motivating me to submit this booklet.
Those of you involved in the 'chicken'-meetings: Abe, Addie, Bas, Esther, Jeroen, Piter, Nancy. Thank you for the fruitful discussions.
Esther, bedankt voor je input en om me af en toe de stal in te nemen om een niet-digitale kip te zien.
Bart en Henk, niet rechtstreeks betrokken bij het project, maar zonder twijfel bijgedragen aan dit boekje. Bart, bedankt voor het beantwoorden van al m'n SAS en ASReml vragen. Henk, bedankt om actief mee te denken.

BIG BIG thank you to all of the (post or present) fellow MSc's and PhD's aka my Wageningen family. Too bad I cannot address every single one of you. This adventure would not have been any fun if it wasn't for you guys. From tea-breaks to road trips, I've enjoyed every single bit of it!

Special thanks to the social girls and the South-American/European bro's.
Social girlssss!!! Coming to work has always been so much fun with the two of you. Thank you for brainstorming with me. There is a reason that both of you are in the acknowledgements of three of my articles. Ewa, thx for the for the get-away's to the ocean, the secret tea-breaks, the parties and the party-aftercare. I would not have survived this PhD without them ;) Naomi, without you the social room would have gotten completely out of hand (aka pink). Thank you for listing to me ramble (in an attempt to understand what I was doing). Pretty sure that if it wasn't for you, this PhD would have taken me even longer ;) It's been fun girl, despite the kicking.
South-American/European bro's. I might have met a bias sample, but I get the impression that I should have been born in a Spanish/Portuguese speaking country! ;) Marcos, I still can't believe we've only met 3-4 years ago. Within no time you managed to get your name on my "best-friends-list". Thank you for always being there for me. Dr. Dr. Juanma, crazy Spanish guy. I'm so glad that you returned to Wageningen. Guess there is such a thing as "second chances". Thank you for being you (crazy I) and allowing me to be me (crazy II). André, normal Spanish (Brazilian-Japanese-Turkish) guy. Thank you for the Monday-distraction, Tuesday-pubquizing and Anyday-friendship. Big G, you always instantly make my day! Thank you for the smiles, hugs (by far the best ones out there) and good advice!

Zodiac-caretakers, reddende engels, dank je om me zo af en toe te hulp te schieten. Het werd enorm geapprecieerd dat er ineens miraculeus een fles cola tevoorschijn kwam als ik triestig naar een lege cola-automaat zat te staren of dat er binnen no-time iemand de printer terug aan de gang kreeg. Merci!!!
Ada, Monique, Lisette, bedankt voor al die goede zorgen! Was altijd fijn om even bij jullie binnen te springen.

Ignacy Misztal, thank you for accepting me as a visiting PhD student. I've learned a lot during my stay at UGA. I really enjoyed working at your department and look back at my time there with nothing but good memories.
Bill Muir, your enthusiasm is infectious! Thank you for the warm welcome, good advice and brainstorm talks. I'm looking forward to the next meet-up, somewhere in world, preferably while drinking a good beer ;)
UGA-people, thank you for accepting me so quickly as part of the group. No better welcome than hearing Dennis say "Are you the new flower?", no better goodbye than hearing Dennis say "I'm going to miss you, sunflower!".

Special thanks to a few of you… Luna, thanks for all the evening distractions. Taking me dancing, organising karaoke or going running. It's been lots of fun ☺ Brenooo, pompo sujo, I really appreciate the help I've got when running the f90-software. Also, thanks for the dinosaur/big lizard and the 7-legged spider. They are truly pieces of art and I'm sure they will make me a millionaire one day. Marianne, I admire you for keeping the pompo sujo in check. Thank you for the warm welcome ;) Dani and Jefferson, you guys have been so nice to me during my visit! Thank you for making me feel at home. The biggest UGA thank you goes out to El. El, you've had a bigger impact on my life than you'll ever realise. Thank you for everything!
Duke, roomie, I'm one lucky girl to have had you as a housemate. Thank you for entertaining me with your helicopter-hawai-twins-guns-boat-crazy gf's-stories, for acknowledging that I'm better at pool than you and for showing me around!

Sarah en Phietje, m'n 2 paranymphen.
Sarah, partner in crime... Dju, kben zo blij dak u heb ☺ Merci dat je er keer op keer voor me bent, no questions asked! Het leven is zo veel makkelijker als je weet dat er iemand is die je opraapt als je tegen de vlakte gaat (letterlijk en figuurlijk). Vriendschap is… Vriendschap is twee kleuren in een pottie ;)
Sophie, misschien zat er tussen ons in wel 700+ km en een uurtje tijdverschil, je bijdrage aan dit boekje is enorm geweest. Grappig hoe wij ooit samen in de auto zaten en het unaniem eens waren dat enkel zotten aan een PhD beginnen. Ik weet niet wat dit nu vandaag over ons beiden zegt ;) Wat heb ik je de afgelopen jaren gemist! Blij dat je terug bent…

En dan rest er mij nog een grote groep Belgen te bedanken: familie, mede-uitgaanders, Berkenbomers uit den ouden tijd, Vikings, Fred, familie Bruyninckx, SJR-ers. Fred, bedankt om me de afgelopen maanden uit mijn vermoeide-stressballetje-modus te proberen halen. Best een heftige eerste relatietest. Blij dat we die ene ochtend aan de praat zijn geraakt, blij dat je vandaag naast me staat. Gino, bedankt vr je steun! Vooral de laatste 2 jaar, toen ik door tijdgebrek ineens veel minder vaak langs kwam, maar je desondanks altijd voor me klaar stond. Het is niet onopgemerkt gebleven. SJR-ers, bedankt voor de afleiding. Ik heb zo genoten van de vrijdagavond-trainingen, samen op pad zondagochtend vroeg (en toch zonder ochtend humeur) ☺ Bedankt voor de steun, vooral toen we op 5 nov 2010 onze kopman verloren. Sven verliezen heeft me doen beseffen hoe kostbaar het leven is… Geen *Carpe Diem* omslag. Wel vlugger zeggen dat je iemand graag hebt, zeggen dat je iemand waardeert, zeggen dat iemand je dag beter maakt.

Mama & papa, ik ben jullie enorm dankbaar voor alle kansen die je me gegeven hebt. Zonder jullie was ik nooit zo ver geraakt. Bedankt voor jullie betrokkenheid, niet alleen naar mij toe, maar ook iedereen om me heen. We worden met z'n allen zwaar verwend! Merci…

Zus, ik weet dat wij water en vuur zijn, maar ik zal je altijd graag zien.

Colophon

The research described in this thesis was financially supported by the Netherlands Organisation for Scientific Research (NWO) and by the Dutch Technology Foundation STW (project number: 10477).

The data used in this thesis were provided by Hendrix Genetics (Boxmeer, The Netherlands).

The cover of this thesis was created by Piotr Sell (www.pcell.it).

This thesis was printed by GVO drukkers & vormgevers B.V. | Ponsen & Looijen, Ede, The Netherlands.