NMBU

# Preface

This master's thesis marks the end of my Master's degree in Animal Sciences with the specialization in Animal Breeding and Genetics at the Norwegian University of Life Sciences (NMBU), Department of Animal and Aquacultural sciences (IHA).

Several people deserve credit for helping me complete this thesis. First I would like to thank my supervisor Prof. Gunnar Klemetsdal at NMBU for his guidance, good discussions and encouragement throughout the process of writing this thesis. I want to thank The Norwegian Association of Sheep and Goat Farmers (NSG) for providing the data material, and for answering my questions. I would also thank my co-supervisor Geir Steinheim for the useful comments, quick responses and support.

A special thank goes to Odin Kirkemoen for supporting me and for spending hours reading through my writing. Finally I would like to thank my uncle, Roger Asbey for helpful input and for revealing mistakes during the writing process.

Norwegian University of Life Sciences

Ås, 1 December 2014

________________________

Siri Storteig Horn

## Abstract

The aim of this study was to evaluate the quality of the data material provided from sheepdog trials in Norway, estimate heritabilities and repeatabilities for the traits included in the trial, as well as make recommendations on how sheepdog trials can best be utilized in breeding of Border Collies. The analyses were based on test results from sheepdog trials carried out in Norway from 1993 to 2012. Originally there were 45 732 records of 3841 Border Collies, but after quality assurance only half were left. The results show that there is little information in the data material. This implies that the evaluation system of the dogs is suboptimal, and/or that a lot of random factors affect the test results. Heritabilities varied between 0.010 and 0.056 with standard errors ranging from 0.010 to 0.023, while repeatabilities ranged from 0.041 to 0.286. Selection based on breeding values would be more precise than the current system which is based on performance results. Nevertheless there would be a need to improve the registration and quality assurance of data to improve the information value of the test results. It is especially important to register the reason behind a zero point score. A further improvement would be to modify the scoring system to prevent the skewed distribution of scores. To increase the value of sheepdog trials in breeding of Border Collies, we recommend adding new, redefined traits to be evaluated in parallel with the current trials. The new traits are based on the Herding Trait Characterization used in Sweden, and also traits from the predatory motor pattern central in Border Collie behaviour.

# Sammendrag

Målet med dette studiet var å evaluere kvaliteten på datamaterialet fra gjeterhundprøver i Norge, estimere arvegrad og gjentaksgrad for momentene i prøven, samt gi anbefalinger for hvordan gjeterhundprøvene best kan brukes i avl på border collie. Analysen er basert på prøveresultater fra norske gjeterhundprøver i perioden 1993 til 2012. I utgangspunktet var det 45 732 observasjoner fra 3841 hunder av rasen border collie, men etter kvalitetssikring av data var bare halvparten igjen. Resultatene viser at det er lite informasjon i datamaterialet. Dette tilsier at evalueringssystemet av hundene ikke er optimalt og/eller at flere tilfeldige faktorer påvirker testresultatene. Arvegradene var mellom 0,010 og 0,056 med standardavvik på 0,010 – 0,023. Gjentaksgradene var også lave, og varierte fra 0,041 til 0,286. Selektering av avlsdyr basert på avlsverdier vil være mer presist enn dagens system som er basert på fenotype. Det vil allikevel være nødvendig å forbedre registreringen av data slik at det blir mulig å utnytte en større del av datamaterialet. Det er spesielt viktig å registrere årsaken til at en hund får 0 poeng på et moment. Poengsystemet bør revurderes for å forhindre skjevfordelingen av poeng. For å øke verdien av gjeterhundprøvene i avl for gjeteregenskapene hos border collie anbefaler vi å legge til nye, redefinerte egenskaper som vurderes parallelt med de nåværende gjeterhundprøvene. Anbefalingene er basert på den svenske ”Arbetsbeskrivning” (HTC), samt elementene i den grunnleggende jaktsekvens som er viktigst for border colliens oppgaver.

# Table of contents

# 1 Introduction

In Norway, sheep are housed in winter, kept grazing on fenced farmland during spring and autumn, and on unfenced mountain or forest pastures during summer. Sheepdogs provide invaluable help to farmers, especially when gathering the sheep in the autumn. Originally, the Norwegian Buhund was assigned this work as it also could serve as a guardian dog protecting against predators (mainly: lynx, wolverine, and wolf). As the predator populations declined in Norway during the 1900's, there was no longer a need for guardian dogs, and the use of dogs in sheep-farming decreased. This changed in the 1950's when Norwegian farmers gained knowledge about the Border Collie (hereafter termed BC) and its role in Great Britain as a specialized herding dog. As a result the import of BCs to Norway began (Drabløs 2003). A significant number of working BCs are still being imported to Norway from the UK and Ireland (Gjerjordet 2013; Størdal 2012), but most dogs are bred domestically.

Today the BC is one of the most popular dog breeds in Norway, both as a working dog and as family companion. Each year 800 to 1000 new BCs are registered in the Norwegian kennel club (NKK) (Norsk Sau og Geit 2014). About 20 % of BCs participate at sheepdog trials (Gjerjordet 2013), and every year over 100 trials are held in Norway. The results from sheepdog trials have been registered in a database since 1994. The Norwegian Association of Sheep and Goat Farmers (NSG) is the breed club for BC in Norway, through an agreement with the NKK. NSGs main goal for the BC is to preserve and develop the breeds herding skills and good health (Norsk Sau og Geit 2014). The sheepdog trials are the most important base for the breeding work. NSG requires that dogs used in breeding have achieved a minimum of 60 points out of 100 on a trial, as well as being healthy and having a good temperament (Norsk Sau og Geit 2014). Litters that don't meet these requirements can be registered in the kennel club, but will not be advertised by the breed club.

Studies of other types of working breeds such as the English Setter (Arvelius & Klemetsdal 2013), Finnish Hounds (Liinamo et al. 1997), Belgian Shepherd (Courreau & Langlois 2005), Finnish Spitz (Karjalainen et al. 1996), German Shepherd and Labrador Retriever (van der Waaij et al. 2008), show that using performance tests to calculate indexes or breeding values would provide great benefits regarding the breeding progress. Arvelius et al. (2013) also states this in a study based on herding behaviour descriptions of BCs in Sweden. In Norway, although the trial is intended to be a tool in genetic evaluation, no genetic parameters are available and no breeding values are estimated. Selection of breeding animals is based on performance results (phenotype). The aim of this study is thus to (1) to evaluate the quality of the data material provided from sheepdog trials in Norway, (2) estimate heritabilities and repeatabilities for the traits included in the sheepdog trial, and (3) to make recommendations on how sheepdog trials can best be utilized in breeding of BCs.

# 2 Material and method

## 2.1 Sheepdog trials

A sheepdog trial is a competition between handler/dog pairs in how they execute a number of herding-elements. The course must contain a pen of 6-8m2 with a swinging gate, three sets of gates, a shedding ring (a marked-out circle of about 40 m in diameter), and a post where the handler stands (Figure 1). At the start, the dog is placed next to the handler with the sheep on the opposite end of the course. The distance (fetch distance) between the dog and the sheep at the start of each run is decided by the judges at each trial.

As in several other European countries, the rules for trials are based on the rules of the International Sheep Dog Society (Hoffmann et al. 2002), but some modifications are made (Gjeterhundrådet 2010; The International Sheep Dog Society 2014). The following describe the current Norwegian rules for trials by the NSG (Gjeterhundrådet 2014).

The first element in every run is *outrun*; the handler sends the dog along the right or left hand side of the course towards the sheep. The dog should run in an arc, ending up behind the sheep without disturbing them. The next element, *lift*, is when the sheep begin to move under the control of the dog. The dog will then bring the sheep in a straight line through gate 1 (Figure 1) and further towards the handler, this is called *fetch* and is considered completed when the flock has circled behind the handler. Next is *Drive*, where the dog must drive the sheep through gate 2 and 3, and then into the pen. In class 1 drive is divided into two elements, *driving in front of handler* and *driving behind handler*. Here the handler must remain at the post until the sheep are through gate 2 and then walk in front of the sheep through the last gate and on to the pen. *Penning* is completed when the

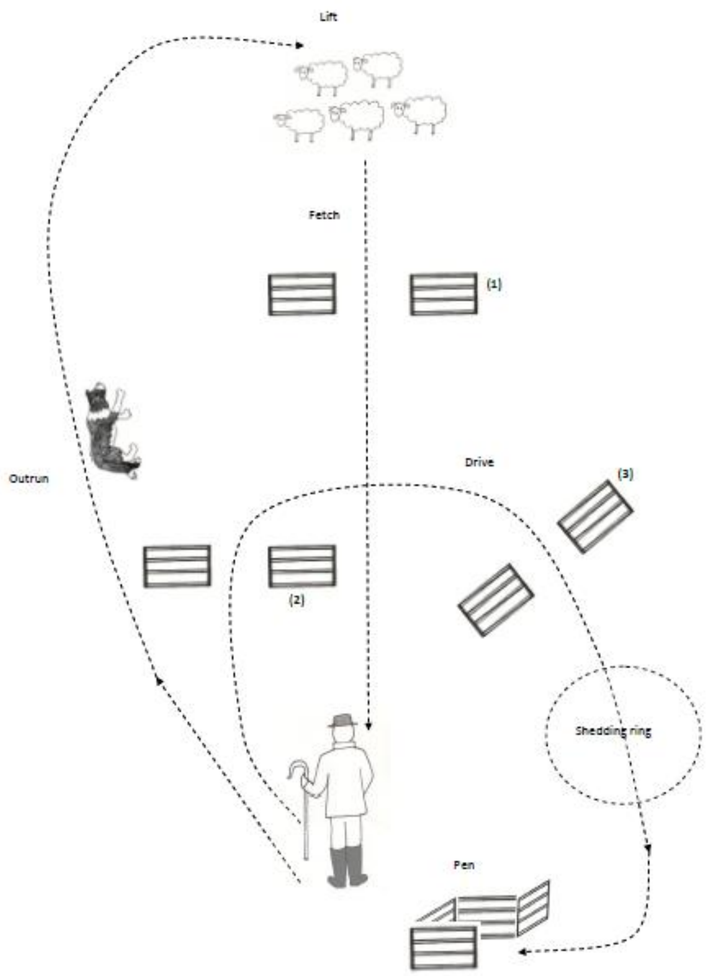


**Figure 1** Drawing of a trial field. The dashed line is demonstrating the run. (1), (2) and (3) refers to the gates the sheep should be driven through.

dog have moved all the sheep into the pen. The elements *Shedding*, *Single* and *Sorting* all take place in the shedding ring, and are different variations of separating specific sheep from the flock or ring.

The sheepdog trials in Norway are divided into three main classes: 1, 2 and 3. There are six different variations of class 3. Thus, a total of 8 different class variants exists (Table 1). The classes differ in which elements are implemented, and the level of difficulty of the elements and the trial in general. Class 1 is the easiest class, while class 3 with double lift is the most difficult class with the most challenging elements.

**Table 1**
Overview of the different classes and the elements that is included in each class, as well as the maximum score for each element. 3a = class 3 with *shedding*, 3b = class 3 with *single*, 3c = class 3 with *shedding* and *single*, 3d = class 3 with *shedding* and *double fetch*, 3e = class 3 with *single* and *double fetch*, 3f = class 3 with *shedding*, *single* and *double fetch*. The table is based on the rules for trials (Gjeterhundrådet 2010).

| Element | Class 1 | Class 2 | Class 3a | Class 3b | Class 3c | Class 3d | Class 3e | Class 3f |
|---|---|---|---|---|---|---|---|---|
| Outrun | 20 | 20 | 20 | 20 | 20 | 40 | 40 | 40 |
| Lift | 10 | 10 | 10 | 10 | 10 | 20 | 20 | 20 |
| Fetch | 25 | 25 | 20 | 20 | 20 | 40 | 40 | 40 |
| Driving | - | 30 | 30 | 30 | 30 | 30 | 30 | 30 |
| Driving in front of handler | 20 | - | - | - | - | - | - | - |
| Driving behind handler | 10 | - | - | - | - | - | - | - |
| Shedding | - | - | 10 | - | 10 | - | - | - |
| Sorting | - | - | - | - | - | 20 | - | 20 |
| Penning | 15 | 15 | 10 | 10 | 10 | 10 | 10 | 10 |
| Single | - | - | - | 10 | 10 | - | 10 | 10 |

Contestants start each element with the maximum score, and the judges then deduct points as they assess the run. Excessive commands from the handler, sheep missing the gates and lack of control of the sheep are considered faults. The "flow" of the run is important for the total score. There are no rules regulating the number of judges on each trial, or how to register the scores from several judges. There can be up to three judges, but most often there is only one. When more than one judge is present, the score from the judges are summarized *or* the judges will agree on a score. Each trial has a time limit set by the judge. If the run is not completed within the time cap, the elements not completed will be scored zero points.

There are several different types of trials. The purpose of the different types is to provide trials of varying difficulty. Larger trials such as regional and national championships will be harder than smaller local trials.

A number of circumstances can result in a score of zero points on an element. Disqualification, withdrawal and not completing within the time cap all lead to 0 points. The different causes are not registered in the database. The least likely alternative is that a dog ends up with 0 points on an element due to poor execution. If an element is carried out it is usually rewarded with at least 1 point (pers. comm. Arne Flatebø).

Factors determining the level of difficulty of each run are class (element composition), type of trial, fetch distance and number of sheep. The Norwegian rules for trials provide guidelines and minimum requirements but no specifications regarding fetch distance and number of sheep.

The scores are not normally distributed (Figure 2). For the elements *outrun*, *lift* and *penning* in both classes the scores are distributed like Figure 2 shows, with a peak at the maximum and minimum scores. *Driving* in class 2, plus *fetch* and *driving in front of handler* in class 1 are more uniformly distributed, while *fetch* in class 2 is the only element with a more normal distribution. All element distributions have a peak at zero points.

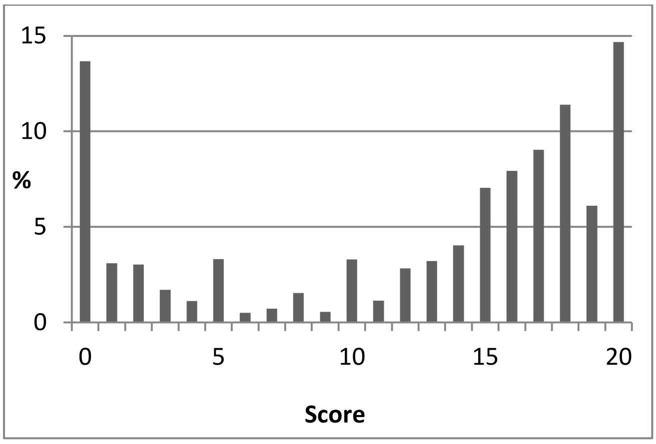


**Figure 2** Distribution of scores for *outrun* class 1.

An important part of the discussion will be whether these trials are the best option regarding measuring the genetics of herding behaviour.

## 2.2 Data material

The data was made available by The Norwegian Association of Sheep and Goat Farmers (NSG), and contained results from 1820 sheepdog trials conducted in Norway in the period from March 1993 to April 2012. Data contained 45 732 records from 3841 border collies. 49.2 % of the records were of male dogs, 50.8 % of females. Most of the dog tested were 3-4 years old.

Data contained information on trial date, location, test-type, test-number, judge, class, fetch distance, owner, handler, kennel, breeder, sex, breed, HD-status, ID, date of birth and parents ID, as well as the trial results including time spent on the trial.

The data was edited to remove obvious errors such as invalid test-type and class, plus illogical fetch distance and dog age. Records missing information on judge, owner or breeder were deleted. Records from contestants that were registered to enter but did not actually enter the trial were identified and deleted. Records with logical errors in the registered scores were also deleted. The remaining data consisted of 26 651 records. These were merged with pedigree data from the Norwegian kennel club (NKK). Records from dogs that didn't exist in the pedigree were deleted. 22641 records remained with an average number of 8.6 records per dog. The pedigree contained information on both parents of all the remaining dogs. The average number of complete generations recorded in the pedigree was 5.9 (Gjerjordet 2013).

Only records from class 1 and 2 was included in the analysis, because all dogs competing in class 3 have already competed in class 1 and/or 2, and the largest amount of records were found in class 2. The rules for trials in class 1 and 2 have not changed during the period from 1994 to 2012, while class 3 rules have changed to some extent (Gjeterhundrådet 2010; Norsk Sau og Geit 1994; Norsk Sau og Geit 1999; Norsk Sau og Geit 2005). The average number of records per dog for class 1 plus 2 was 6.3, with a total of 16447 records.

Season of the year (winter, spring, summer or autumn) was calculated from the date of trial. Event number was made to identify dogs being tested under the same conditions (the same day, location and class). The date of trial together with the dogs' birth date was used to calculate the age at trial. Age was classified into three groups where group 1 contained dogs younger than two years, group 2 dogs from 2 to 4 years, and group 3 dogs older than 4 years. Fetch distance ranged from 48 – 800 m, and was divided into 3 groups: less than 100m, 101-150m, and more than 151m (because very few dogs were tested over 150 m in class 1 and 2). One of the test-types only had 12 records in class 1 and 2 and was excluded from the analysis.

Any zero points obtained in trial were excluded from the analysis because it was not possible to know whether the zero points was due to the performance on the trial or one of the many other ways to achieve zero points. For class 1 (5581 records) 15 % of the results for *outrun*, *lift*, *fetch* and *driving behind handler* were scored zero points and excluded from the analysis. The percentage was 22 for *penning* and 37 for *driving in front of handler*. In class 2 (10866 records) about 10 % of the *outrun*, *lift* and *fetch* results, 16 % of *driving* and 21 % of *penning* results were scored zero points and excluded.

## 2.3 Statistics

Initially a univariate analysis was carried out including data from class 1 only, treating the scores of the behavioural traits as quantitative traits. The following linear animal model was assumed for each element of the trial:

$$\mathbf{Y_{ijklmnopqrst} = \mu + sex_j + year_k + age_l + type_m + season_n + distance_o + judge_p + event_q + a_r + pe_s + e_{ijklmnopqrs}}$$

Where $Y_{ijklmnopqrst}$ = a score on the performance of an element of the trial, μ = overall mean, $sex_j$ = fixed effect of the $j^{th}$ sex (j=1,2), $year_k$ = fixed effect of the $k^{th}$ year (k=1994-2012), $age_l$ = fixed effect of the $l^{th}$ age group (l=1-3), $type_m$ = fixed effect of the $m^{th}$ test-type (m=1-5), $season_n$ = fixed effect of the $n^{th}$ testing season (n=1-4), $distance_o$ = fixed effect of the $o^{th}$ fetching distance group (o=1-3), $judge_p$ = random effect of the $p^{th}$ judge (judge=1-119), $event_q$ = random effect of the $q^{th}$ event-number/unique event (ID) (q=1-2299), $a_r$ = random additive genetic effect of the $r^{th}$ animal, $pe_s$ = random permanent environmental effect associated with the $s^{th}$ animal, and $e_{ijklmnopqrs}$ = random residual effect.

The model had the following variance-covariance structure:

$$V\begin{bmatrix} a \\ pe \\ judge \\ event \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_a^2 & 0 & 0 & 0 & 0 \\ 0 & I_{pe}\sigma_{pe}^2 & 0 & 0 & 0 \\ 0 & 0 & I_{judge}\sigma_{judge}^2 & 0 & 0 \\ 0 & 0 & 0 & I_{event}\sigma_{event}^2 & 0 \\ 0 & 0 & 0 & 0 & I_e\sigma_e^2 \end{bmatrix}$$

A is the additive relationship matrix and I are identity matrices.

Judge and event were considered random effects because these sub-classes had few records, and to decide whether to include them models with or without judge and event were compared for their -2ln likelihoods. First, a model without both judge and event was run, then with judge, and lastly with both judge and event.

Finally, four bivariate analyses were carried out, one for each of the elements that was included in both class 1 and 2 (*outrun*, *lift*, *fetch* and *penning*). The same class 1 data as in the univariate analysis was used. Data from class 2 was selected based on the same principles as for class 1. The same independent model terms were used as in the univariate analysis.

The expectations of random effects were all zero with the following distributions:

$$V\left(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}\right) = \begin{bmatrix} \sigma^2_{a_1} & \sigma_{a_1\,a_2} \\ sym. & \sigma^2_{a_2} \end{bmatrix}$$

$$V\left(\begin{bmatrix} pe_1 \\ pe_2 \end{bmatrix}\right) = \begin{bmatrix} I\sigma^2_{pe_1} & I\sigma_{pe_1\,pe_2} \\ sym. & I\sigma^2_{pe_2} \end{bmatrix}$$

$$V\left(\begin{bmatrix} judge_1 \\ judge_2 \end{bmatrix}\right) = \begin{bmatrix} I\sigma^2_{judge_1} & I\sigma_{judge_1\,judge_2} \\ sym. & I\sigma^2_{judge_2} \end{bmatrix}$$

$$V\left(\begin{bmatrix} event_1 \\ event_2 \end{bmatrix}\right) = \begin{bmatrix} I\sigma^2_{event_1} & I\sigma_{event_1\,event_2} \\ sym. & I\sigma^2_{event_2} \end{bmatrix}$$

$$V\left(\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}\right) = \begin{bmatrix} I\sigma^2_{e_1} & 0 \\ sym. & I\sigma^2_{e_2} \end{bmatrix}$$

No convergence could be achieved for the bivariate analysis of *penning*. The analysis converged when the variance for *pe* was fixed at 0.435 as estimated for $\sigma^2_{pe}$ in the univariate analysis.

All analyses were performed using ASREML software (Gilmour et al. 2002).

# 3 Results

Table 2 shows that all effects except distance were significant (p-value < 0.05) for at least one of the elements of the trial. A strong significance (p-value < 0.001) was found for year on *penning* and for age on *outrun*.

**Table 2**
Levels of significance for the effects of sex, year, age, test-type, season and distance from the univariate analysis of class 1 data.

| | Element | | | | | |
|---|---|---|---|---|---|---|
| Environmental factor | Outrun | Lift | Fetch | Driving in front | Driving behind | Penning |
| Sex | 0.894 | 0.503 | 0.236 | 0.492 | * | 0.593 |
| Year | 0.085 | 0.513 | 0.415 | 0.250 | * | *** |
| Age | *** | 0.492 | 0.680 | ** | ** | ** |
| Test-type | 0.143 | * | * | 0.648 | 0.304 | 0.988 |
| Season | 0.184 | 0.809 | 0.534 | 0.579 | 0.214 | *** |
| Distance | 0.976 | 0.647 | 0.592 | 0.081 | 0.708 | 0.777 |

* P < 0.05
** P < 0.01
*** P < 0.001

When applying likelihood-ratio-test the improvement was significant for the random effects of both judge and event on all elements of the trial (Table 3).

**Table 3**
$\chi2$ –values for the likelihood-ratio-test of two models, first with judge as a random effect in the model, and then adding event to the model. $\chi2$ -values > 7.88 are significant (P<0.05) for 1 degree of freedom.

| Element | Judge | Event \| Judge |
|---|---|---|
| Outrun | 24.8 | 13.4 |
| Lift | 46.5 | 57.3 |
| Fetch | 128.2 | 125.0 |
| Driving in front | 38.6 | 42.9 |
| Driving behind | 161.8 | 212.1 |
| Penning | 49.8 | 37.2 |

As seen in table 4, all heritabilities except *penning* in class 1 were significantly different from zero. Heritabilities ranged from 0.010 to 0.056, *penning* having the lowest and *driving in front of handler* having the highest heritability (both in class 1). Estimates of repeatability were low, ranging from 0.041 for *penning* in class 1 up to 0.286 for *driving* in class 2 (Table 4).

**Table 4**
Variance components ($\sigma^2$) for additive genetic (a), permanent environment (pe), judge, event and the remaining residual (e) effects for the elements of class 1 and 2. Plus calculated values for heritability ($h^2$) and repeatability (r) with standard errors (SE).

| | Element | $\sigma^2_a$ | $\sigma^2_{pe}$ | $\sigma^2_{judge}$ | $\sigma^2_{event}$ | $\sigma^2_e$ | $h^2$ (SE) | r (SE) |
|---|---|---|---|---|---|---|---|---|
| CLASS 1 | Outrun | 1.510 | 5.118 | 0.517 | 1.003 | 23.714 | 0.047 (0.018) | 0.208 (0.019) |
| | Lift | 0.105 | 0.559 | 0.138 | 0.381 | 4.084 | 0.020 (0.013) | 0.126 (0.018) |
| | Fetch | 1.656 | 3.106 | 1.155 | 3.981 | 26.391 | 0.046 (0.017) | 0.131 (0.017) |
| | Driving in front | 1.553 | 2.896 | 0.871 | 2.176 | 20.206 | 0.056 (0.023) | 0.161 (0.022) |
| | Driving behind | 0.094 | 0.415 | 0.409 | 1.077 | 4.726 | 0.014 (0.012) | 0.076 (0.015) |
| | Penning | 0.147 | 0.435 | 0.339 | 0.942 | 12.429 | 0.010 (0.012) | 0.041 (0.016) |
| CLASS 2 | Outrun | 0.818 | 2.941 | 0.301 | 0.569 | 12.972 | 0.047 (0.020) | 0.214 (0.014) |
| | Lift | 0.136 | 0.387 | 0.032 | 0.213 | 2.741 | 0.039 (0.015) | 0.149 (0.012) |
| | Fetch | 0.987 | 6.292 | 0.971 | 3.315 | 17.131 | 0.034 (0.016) | 0.254 (0.014) |
| | Driving | 2.238 | 11.294 | 3.090 | 4.686 | 26.048 | 0.047 (0.021) | 0.286 (0.016) |
| | Penning | 0.210 | - | 0.095 | 1.073 | 7.605 | 0.023 (0.010) | 0.055 (0.008) |

$h^2 = \sigma^2_a / \sigma^2_a + \sigma^2_{pe} + \sigma^2_e + \sigma^2_{event} + \sigma^2_{judge}$
$r = \sigma^2_a + \sigma^2_{pe} / \sigma^2_a + \sigma^2_{pe} + \sigma^2_e + \sigma^2_{event} + \sigma^2_{judge}$

The results from the bivariate analysis show that additive genetic correlations for the four elements were positive, but had considerable standard errors (Table 5). They were low for *outrun* and *lift*, and higher for *fetch* and *penning*, indicating that the elements common to the two classes cannot be considered the same traits. There were rather high correlations for the effect of permanent environment, which implies that dogs that get high scores in class 1 do so in class 2 as well. The judge correlations are high, meaning that each judge seem to judge the elements similarly in both classes.

**Table 5**
Estimates of additive genetic ($r_g$), permanent environment ($r_{pe}$) and judge ($r_j$) correlations and their standard errors (SE) obtained in bivariate analyses utilizing data from class 1 and 2.

| Element | $r_g$ (SE) | $r_{pe}$ (SE) | $r_j$ (SE) |
|---|---|---|---|
| Outrun | 0.410 (0.259) | 0.741 (0.100) | 0.845 (0.140) |
| Lift | 0.366 (0.354) | 0.836 (0.128) | 0.725 (0.210) |
| Fetch | 0.787 (0.246) | 0.808 (0.118) | 0.822 (0.128) |
| Penning | 0.953 (0.306) | - ( - ) | 0.810 (0.224) |

# 4 Discussion

In general, low repeatability and heritability estimates were found, intervals ranging from 0.041 to 0.286, and from 0.010 to 0.056, respectively. It is relevant to compare the results of this study to studies of hunting behaviour in dogs, as hunting and herding behaviour is constitutionally alike (Coppinger 1995). The heritabilities found in this study are lower than most studies of hunting behaviour in dogs (Arvelius & Klemetsdal 2013; Brenøe et al. 2002; Karjalainen et al. 1996; Liinamo et al. 1997; Lindberg et al. 2004). In a genetic analysis of herding behaviour in BCs, Arvelius et al. (2013) found very high heritabilities, based on Herding Trait Characterization (HTC), not competition results. The most comparable study is by Hoffmann et al. (2003) who analyzed results of sheepdog trials and found heritabilities between 0.001 and 0.13, which is similar to the results of the present study.

The low repeatability implies that the evaluation system of the dogs is suboptimal, and/or that a lot of random factors affect the test results. The traits the judges evaluate are open to interpretation, one element can consist of more than one trait, and it is possible that two dogs that get the same score on an element can have executed the task in different manners.  Further, the distribution of scores isn’t the most favourable (Figure 2), a normal distribution would be optimal statistically, and it would be an indication of the tests being more objectively judged. Using the whole scoring-scale would be beneficial as it would increase variation and information in the data material. The dogs are being judged on a scale from bad to good, which doesn’t give any information on why the dogs are considered bad or good. This scale also makes the judges more likely to give a score to the upper end of the scale.

The heritabilities are too low to achieve substantial genetic progress when using a phenotype-based selection; however, if selection of breeding animals is based on estimated breeding values (EBVs) with use of repeated records, the accuracy will increase. In this data material there was an average of 6.3 observations on each animal, which is expected to increase the accuracy of the EBVs 2-3 times (Danell 1984), with proportionally the same effect on genetic response. By analyzing each trait in class 1 and 2 bivariately we obtained a maximum of 11 EBVs per animal. The analysis did not utilize information between traits within each class, or between different traits in the two classes. The latter was not done because we had to fix parameters already when estimating the correlations between the equivalent elements in the two classes. The fact that no convergence could be achieved for the bivariate analysis of *penning* implies that despite the large amount of data, there is not a lot of information in the data material.

A considerable amount of the data material could not be used in the analysis because the quality of the data was poor. The original data material consisted of 45 732 records, but after quality assurance only half (22641 records) were left. When we deleted the records from class 3 we were left with 16447 records. Additionally, records with zero point scores had to be excluded from the analysis, which constituted 10 to 37 % of the remaining records. There is definitely a need to reduce the loss of data, especially records with zero point scores. The registration of results from each trial is done by the local arranger. Therefore it is especially important that the registration scheme makes it impossible to register illogical information (e.g.: a class that doesn't exist), and that all information has to be submitted. For instance: handler is a factor that would be desirable to include in the model, but because the name of handlers are registered manually each run, rather than giving each handler a unique number that can be cross-checked with a registry, it was impossible to use in the analyses. It is also problematic that the rules for trials make it possible to adjust number of sheep and judges, as well as time limit at each trial, and that this is not registered anywhere. It may be that in reality there are no big variations between trials, but this isn't possible to confirm by looking at the data material. The consequence of not registering the number of judges, and the way the judges register the scores, is that trials with more than one judge had to be excluded from the analysis. A vital improvement would be to register the reason behind a zero point score. It is important to distinguish between withdrawal, exceeding the time limit and disqualification, especially if the disqualification was due to the dog biting the sheep, as this behaviour is not desirable in the BC, and is likely to be heritable (Coppinger 1995).

In dogs, herding behaviour patterns are considered to be strictly heritable (Coppinger 1995), as differences in herding behaviour is apparent between the different herding dogs (headers, heelers and catch dogs). Also, it is known that breeds lose their herding abilities when these are no longer selected for, as has happened to the German Shepherd dog (Houpt 2001). Herding behaviour is in essence predatory behaviour to which sheep are especially sensitive (Houpt 2001; Spady & Ostrander 2008). The basic predatory motor pattern is described in similar manners in both Scott and Fuller (1965) and Coppinger and Coppinger (2001), and can be presented as a sequence of: *orient* → *eye-stalk* → *chase* → *grab-bite* → *kill-bite* → *dissect* → *consume* (Coppinger & Coppinger 2001). This is not behaviour that dogs learn, it is instinctual, but can be modified by artificial selection, thus considered heritable (Coppinger & Coppinger 2001; Scott & Fuller 1965). Centuries of artificial selection particularly for changing this pattern have only resulted in changes in intensity and display of the different elements, not in the basic structure (Coppinger & Coppinger 2001; Scott & Fuller 1965). Comparing the behaviour of the BC to the predatory motor pattern, the *eye-stalk* and *chase* have been intensified, and the *grab*- and *kill-bite* diminished (Coppinger 1995). It is apparent that the behaviour of different hunting dogs (e.g.: Pointers, Retrievers) can also be described through modification of this pattern. This suggests that one should breed for traits that could be defined into the predatory motor pattern, and that one should search for alternative tests that would give better results than the current sheepdog trials.

An alternative to the sheepdog trials would be to adopt the HTC used in Sweden, where heritabilities are found to be high (0.14-0.50) for version 1 (Arvelius et al. 2013). Both countries could benefit from this as shown by Arvelius and Klemetsdal (2013) since the genetic material is shared (Norsk Sau og Geit 2014). The HTC is a description of the behaviour of the dogs, rather than a competition. The intensity of the expression of the measured traits is evaluated on an objective scale, and the test takes place over a longer period of time, as an evaluation during sheepdog training. Traits such as *power* and *affability towards humans* are evaluated in addition to herding behaviour. *Outrun* and *lift* are traits that are being judged both in the HTC and the Norwegian sheepdog trials, but the estimated heritabilities are substantially higher in the HTC. A crucial factor probably contributing to the big difference in heritability estimates is the definition of traits and scoring scale. The traits in the HTC are specific and easy to judge, while the traits in the sheepdog trials are complexly defined in the rules for trials, and a lot of factors influence the score (Gjeterhundrådet 2010; Gjeterhundrådet 2014; Norsk Sau og Geit 1994; Norsk Sau og Geit 1999; Norsk Sau og Geit 2005).

It might be difficult to get people to participate at a new test in addition to the sheepdog trials, especially one that is not a competition. Therefore, it would be an advantage to implement new traits that can be judged in conjunction with the current sheepdog trials. The following proposal of new trait definitions is based on traits from the predatory motor pattern and traits from the HTC. Traits that we recommend adopting from the HTC are traits with high heritability and that are possible to evaluate during a sheepdog trial. These are *outrun*, *lift*, *working distance* and how frequently the dog *loses sheep*. Traits from the predatory motor pattern that are most important for the BC are; intensity and shape of *chase*; intensity of *eye-stalk*; if and how often the dog *bite sheep*; and finally, the *threshold* between the motor patterns, as this is a crucial factor determining the trainability of the BC according to Coppinger and Coppinger (2001). In total this is 9 traits that can be judged parallel to the current trials. All new traits except *outrun* and *lift* will be judged as an evaluation of the behaviour during the whole trial. It is important that the test is standardized, has a precise, neutral and objective scoring scale, and provide sufficient amount of data and information (Arvelius et al. 2013; Diederich & Giffroy 2006).

20 years of data material from sheepdog trials in Norway are available. It would be an advantage to be able to utilize these in breeding of BC. Implementing EBVs will give greater accuracy than phenotype selection. The majority of the elements of the sheepdog trials are comprised of *eye-stalk* and *chase* behaviour from the predatory sequence. Therefore, we would suggest selection based on EBVs for *lift*, as this element is most similar to, and limited to only one of the predatory motor patterns; *eye-stalk*. *Outrun* can also be a candidate for EBV selection, because it might be equivalent to *orient* of the predatory motor pattern, though less obvious. To increase the accuracy and improve the value of sheepdog trials in breeding there is definitely a need to redefine traits as described above.

## 5 Conclusion

Low heritabilities and repeatabilities indicate that the traits evaluated are influenced by several environmental factors that could not be corrected for in the analysis, and/or that there are several behavioural traits included in each of the elements of the trial. Thus, selecting breeding animals based on performance on sheepdog trials is unreliable. Selection based on breeding values would be more precise, using all genetic links and correcting for systematic effects. Still, there would be a need to improve the registration and quality assurance of data to improve the information value of the test results. It is necessary to make sure all information has to be submitted and that it impossible to register illogical information. It is especially important to register the reason behind a zero point score. A further improvement would be to modify the scoring system to prevent the skewed distribution of scores, thus increase variation and information in the data material. The traits in the sheepdog trial are open to interpretation, which makes it possible that two dogs that get the same score on an element can have executed the task in different manners. Plus, one element can consist of more than one trait. Dogs are being evaluated on a scale from bad to good, rather than on a descriptive scale. The results implies that the current sheepdog trials are not optimal as a basis for selection of breeding animals, and that one should search for alternative tests that would give better results. A more advantageous test evaluates the dogs on an objective, neutral scale, with precisely defined traits. A good example of this is the Swedish HTC. Given that herding behaviour is in essence predatory behaviour with presumably high heritability, one should breed for traits that could be defined into the predatory motor pattern. Therefore, we recommend adding new, redefined traits based on traits from the HTC and the predatory motor pattern to be evaluated parallel to the current trials.

# 6 References


Arvelius, P. & Klemetsdal, G. (2013). How Swedish breeders can substantially increase the genetic gain for the English Setter's hunting traits. *J Anim Breed Genet*, 130 (2): 142-53.

Arvelius, P., Malm, S., Svartberg, K. & Strandberg, E. (2013). Measuring herding behavior in Border collie-effect of protocol structure on usefulness for selection. *Journal of Veterinary Behavior-Clinical Applications and Research*, 8 (1): 9-18.

Brenøe, U. T., Larsgarda, A. G., Johannessen, K.-R. & Uldal, S. H. (2002). Estimates of genetic parameters for hunting performance traits in three breeds of gun hunting dogs in Norway. *Applied Animal Behaviour Science*, 77 (3): 209–215.

Coppinger, R. & Coppinger, L. (2001). *Dogs: A Startling New Understanding of Canine Origin, Behavior & Evolution*. New York, USA: Scribner.

Coppinger, R., Schneider, R. (1995). Evolution of working dogs. In Serpell, J. (ed.) *The Domestic Dog: Its Evolution, Behaviour and Interactions with People*, pp. 21-50. Cambridge, UK: Cambridge University Press.

Courreau, J.-F. & Langlois, B. (2005). Genetic parameters and environmental effects which characterise the defence ability of the Belgian shepherd dog. *Applied Animal Behaviour Science*, 91: 233-245.

Danell, Ö. (1984). Kvantitiativ genetik och allmän husdjursförädling. Chapter 10. p. 11.

Diederich, C. & Giffroy, J.-M. (2006). Behavioural testing in dogs: A review of methodology in search for standardisation. *Applied Animal Behaviour Science*, 97 (1): 51-72.

Drabløs, D. (2003). *Gjetartradisjonar i Noreg: sosiale, biologiske og teknologiske utviklingsdrag*. Høvik: Kolofon. 198 pp.

Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J. & Thompson, R. (2002). *ASReml User Guide Release 1.0.VSN Int. Ltd.* Hemel Hempstead, UK.

Gjerjordet, H. F. (2013). *Pedigree analysis on border collie data in Norway*. Ås, Norway: Norwegian University of Life Sciences, Department of Animal and Aquacultural Sciences. 26 pp.

Gjeterhundrådet. (2010). *Regler for gjeterhundprøver*. Groseth, A., Kaasa, O., Skjærvold, H., Arnesen, A. & Størdal, S. R. (eds). Ås: Norsk Sau og Geit. p. 17.

Gjeterhundrådet. (2014). *Regler for gjeterhundprøver*. Loftsgarden, A., Bjørdal, O. B., Sand, A. C., Nymo, K., Seilen, A. & Kjelstrup, C. (eds). Ås: Norsk Sau og Geit. p. 35.

Hoffmann, U., Hamann, H. & Distl, O. (2002). Genetic analysis of traits of the working test for Border Collies. *Zuchtungskunde*, 74 (3): 213-231.

Hoffmann, U., Hamann, H. & Distl, O. (2003). Genetic analysis of traits of the working test for herding dogs 1st communication: Performance traits. *Berliner Und Munchener Tierarztliche Wochenschrift*, 116 (3): 81-89.

Houpt, K. A., Willis, M.B. (2001). Genetics of behaviour. In Ruvinsky, A. & Sampson, J. (eds) *The Genetics of the Dog*, pp. 371–400. Wallingford, Oxon, UK: CABI.

Karjalainen, L., Ojala, M. & Vilva, V. (1996). Environmental effects and genetic parameters for measurements of hunting performance in the Finnish Spitz. *Journal of Animal Breeding and Genetics*, 113 (1-6): 525-534.

Liinamo, A. E., Karjalainen, L., Ojala, M. & Vilva, V. (1997). Estimates of genetic parameters and environmental effects for measures of hunting performance in Finnish hounds. *J Anim Sci*, 75 (3): 622-9.

Lindberg, S., Strandberg, E. & Swenson, L. (2004). Genetic analysis of hunting behaviour in Swedish Flatcoated Retrievers. *Applied Animal Behaviour Science*, 88 (3-4): 289-298.

Norsk Sau og Geit. (1994). *Reglar for gjetarhundprøver*. Oslo. p. 7.

Norsk Sau og Geit. (1999). *Reglar for gjetarhundprøver*. Oslo. p. 9.

Norsk Sau og Geit. (2005). *Reglar for gjetarhundprøver*. Oslo. p. 7.

Norsk Sau og Geit. (2014). *Rasespesifikk Avlsstrategi for Border Collie*. Available at: http://www.nsg.no/getfile.php/_NSG-PDF-filer/Gjeterhund/NKK/RAS_28102014.pdf (accessed: 24.11.2014).

Scott, J. P. & Fuller, J. L. (1965). *Genetics and the Social Behavior of the Dog*: University of Chicago Press.

Spady, T. C. & Ostrander, E. A. (2008). Canine behavioral genetics: Pointing out the phenotypes and herding up the genes. *American Journal of Human Genetics*, 82 (1): 10-18.

Størdal, S. R. (2012). Unik gjeterhund og en etterspurt sportshund. *Sau og Geit*, 65 (2): 50-51.

The International Sheep Dog Society. (2014). *Rules for Trials*. In Sheen, J. (ed.). Bedford, UK. Available at: http://www.isds.org.uk/download/clientfiles/files/Rules%20For%20Trials%20(2014)(1).pdf (accessed: 24.11.2014).

van der Waaij, E. H., Wilsson, E. & Strandberg, E. (2008). Genetic analysis of results of a Swedish behavior test on German Shepherd Dogs and Labrador Retrievers. *J Anim Sci*, 86 (11): 2853-61.

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no