2017

An investigation into the usefulness of the British Equestrian Federation Futurity programme

Lauren Brown

RESM

An investigation into the usefulness of the British Equestrian Federation Futurity programme

Brown, Lauren

Award date:
2017

Awarding institution:
University of Plymouth

An investigation into the usefulness of the British Equestrian Federation Futurity programme

by
Lauren Brown

A thesis submitted to Plymouth University in fulfilment for the degree of

Research Masters in Equitation Science

April 2017

Abstract.

An investigation into the usefulness of the British equestrian Federation Futurity programme

By Lauren Brown

Sport horse genetic evaluations are used throughout Europe to inform breeding practices. In England specifically, British Equestrian Federation Futurity evaluations are used to assess horse potential and to inform British breeding. Futurity premium scores are allocated to represent performance potential based on horse's component traits (primarily conformation and locomotion). This study demonstrated that over half (61/106) questionnaire respondents believe that Futurity premium scores influence horse training decisions and monetary value. This is highly relevant to the industry as premium scores can therefore have economic and welfare implications. Therefore horse premiums must be reliable and indicative of future competition performance. Retrospective Futurity premiums (n=566) were investigated for reliability using the MiniTab™ v17 statistical package. Regression analysis of premium scores versus test variables demonstrated that whilst horse age and colour significantly correlated with Futurity premiums (P=0.017 and P=0.027 respectively); sex and test location did not (P >0.05). Regression analysis of components of horse competition records (lifetime points, placings, winnings, percentage scores and penalty points) versus Futurity premiums demonstrated limited correlations. British Eventing penalty point scores correlated with the majority of Futurity scores, BD percentage scores correlated with one and British Showjumping, none. British Eventing penalty point scores appear the most appropriate measure of performance as this method evaluates each phase of the individual's competition, unlike points/placings. Futurity component scores demonstrated mixed results. A high Futurity score did not necessarily predict a high performance score. A key issue with the data was the fact that the horses examined were relatively young. Average horse age was 7 years and therefore these horses may not have been old enough to have developed their abilities towards their mature potential, limiting competition results therefore their records and consequently affecting analysis outcomes. Furthermore, industry practitioners have highlighted judge subjectivity as a limitation, however results suggest that test location (representing the judging panel) is not influential in scoring. As the Futurity develops, a larger dataset of older horses will become available which will provide further insight into the tests usefulness.

Keywords: BEF Futurity, sport horse, performance, genetic evaluation and breeding.

Contents **Page**

List of tables, figures and appendices.

List of tables **Page**

1. Tables

List of figures

2. Figures

List of appendices

3. Appendices

AUTHOR'S ACKNOWLEGMENTS

First, thank you to the British equestrian Federation for providing me with the raw data for the study. Second, thank you to my supervisors Hayley Randle and Peter McGregor for supporting me along the way to completing this research. Thank you to Emma Davies for helping when needed and finally I am grateful to the European Social Fund for giving me the means to enrol and complete my masters.

AUTHOR'S DECLARATION

At no time during the registration for the degree of Research Masters (Equitation Science) has the author been registered for any other University award without prior agreement of the Graduate Committee.

Work submitted for this research degree at the Plymouth University has not formed part of any other degree either at Plymouth University or at another establishment.

Presentation and Conferences Attended: The British Society of Animal Science (April 2015 attendance) and Breeding for Gold (March 2016 presentation).

External Contacts: Jan Rogers (Head of Development at the British Equestrian Federation).

Word count of main body of thesis: 24,595

Signed: *Lauren Brown* **Date:** 11.04.17

Chapter 1. A review of the literature examining sport horse phenotypic evaluations.

The purpose of selective breeding is to increase or decrease certain genetic biological traits within an animal population (Villemereuil 2012). A biological trait is a physiological trait of an individual which is controlled by the individual's genotype (collection of genes) and are identifiable by their phenotypes. For instance, observable hair colour is the phenotype of the genotypic DNA (Suzuki and Griffiths 2000). Selective breeding aims to promote genetic gain through the breeding of desirable genotypic DNA which can be identified by the corresponding phenotype. Desirable biological traits can then be inherited by the proceeding generations (Gjedrem and Baranski 2009). The success of selective breeding relies on the heritability of the desired genotype. Heritability is a statistic used in selective breeding which estimates how much variation in phenotypic traits of a population is the result of genetic variation among individuals in that population. Heritability is estimated by comparing individual phenotypic variation among differently related groups, such as offspring of the same parentage (Falconer and Mackay 1995). The heritability value describes the likelihood of a biological trait being transmitted within a population from generation to generation (Villemereuil 2012). Heritability increases when genetics are contributing more to the phenotypic trait being measured and/or because non-genetic (environment) factors are contributing less. Genetics effects of dominant/recessive genes and gene linkage can skew inheritance predictions. In addition, environment factors such as malnutrition can create variability in the physical appearance of phenotypes. Nevertheless, Villemereuil (2012) suggests that positive heritability rates can be predicted from well-designed and controlled selective breeding programs by accounting for the genetic influence and estimating the effects of environmental variance. A criterion for selecting animals for breeding is therefore a high heritability of desired biological traits. A highly heritable trait is a trait rich in genetic information, increasing the accuracy of improving upon the next generation produced and consequently increasing the odds of genetic gain (Pryce 2011). Animal phenotypic evaluations are used to assess the phenotype occurrence of traits to inform breeding practices. Selection criteria of phenotypic evaluations depends on the traits which are being assessed.

Phenotypic evaluations in the Livestock industry

Phenotypic evaluations are used within the livestock industry to predominantly identify highly heritable and desired biological traits in order to improve the productivity of the following generations. Livestock farmers have been selecting breeding animals on the basis of their phenotypes for centuries. Phenotypic evaluations as they are known today were first introduced some 20-30 years ago when Britain first published a breeding reference book of cattle records (Epstein & Mason 1984). Today, animal phenotypic evaluations are informed

by assessing the pedigree and performance data of the individual. Livestock pedigree and performance data (also known as phenotype records) can be used to estimate the breeding value (EBV) of an individual animal, this is most commonly undertaken by using a statistical model known as best linear unbiased selection (BLUP) (Goddard et al. 2010). Weller et al. (2015) reviewed the benefits of phenotypic evaluations based on EBV's calculated from pedigree and performance data to improve dairy cattle productivity and demonstrated that through breeding livestock based on their individual milk yield profiles and parentage information, the overall quantity of milk produced was improved. Similarly, within the beef cattle industry, production goals are to breed larger animals in order to obtain greater meat yield per carcass, therefore improving food conversion efficiency and reducing rearing cost per kilogram sold. Hayes et al. (2013) demonstrated the benefit of using the weight measures of beef cattle to calculate EBV's for genetic gain. EBV's are also used in the sheep farming for the selection and breeding of production related traits such as wool quality, milk yield, and fertility (Carta et al. 2009 and Safari et al. 2007).

Whilst phenotypic evaluations in the livestock industry clearly contribute to increase in animal productivity, some concerns are present. Taberlet et al. (2008) argue that intensive phenotypic selection programs in livestock farming can lead to a decrease in population genetic diversity and subsequently decreases in animal health. Genetic diversity is important for the long term health of groups of animal as first, population groups become less susceptible to inherited disorders and second, genetic variety makes breed advances possible. Taberlet et al. (2008) maintain that selective breeding programs must be carefully managed to prevent genetic uniformity, through decreasing inbreeding and introducing unrelated animals into existing animal breeding groups. Genetic diversity can be maintain through awareness of animal pedigree history. The Department for Environment, Food and Rural Affairs (Defra) support selective breeding by funding BASCO, a collective hub of livestock pedigree and performance data for breeders to implement into their programmes to obtain breeding goals (Coffey 2007).

Phenotypic evaluations in the working dog industry

Phenotypic assessments are used in a range of working dog industries to assess dog performance ability in order to achieve breeding goals. For example in guide dog breeding, the objective is to breed dogs of sound health and temperament for training (Guide Dogs 2015). Trybocka (2010) investigated the suitability of character testing of trainee guide dogs to assess their performance ability by comparing early character testing results with later in life performance. Trybocka (2010) concluded that character testing was a true reflection of a dog's suitability to guide dog work as moderate trait heritabilities were predicted from their behaviour, suggesting that some genetic merit can be estimated. However the study stated

more data is needed to predict significant results for confident use in the industry and therefore the results must be interpreted with caution. In addition to phenotypic evaluations of guide dogs, military working dogs are additionally performance tested. Phenotypic evaluations of military working dogs are used to achieve the breeding goal of producing dogs which are trained to respond consistently to reactive stimuli in any given situation (Military Police 1993). Sinn et al. (2010) demonstrated that behaviour testing in military working dogs can predict positive, but low trait heritabilities and emphasises that a greater understanding is needed of which genotypic traits relate to the desired tested phenotypic traits in order to predict more accurate heritability estimations of genetic merit. Sinn et al. (2010) demonstrated the importance of understanding the relation between genotypic traits and their phenotypic appearance in evaluations.

Phenotypic evaluations in the Equine industry

It is clear from existing published research that phenotypic evaluations are used in industry specific breeding programmes to help achieve required breeding goals. In the Equine industry, breeding goals are achieved through selective breeding programs, informed by phenotypic evaluations, similar to in the livestock and working dog industries. The breeding goals of the equestrian industry vary depending on the work type a horse is being bred for. Heavy horse breeds, such as the British Percheron or Suffolk horse were originally bred for their capabilities in pulling heavy load carriages, therefore these breeds were selected based on their phenotypic display of strength and a calm temperament (The British Percheron Society 2009 and Suffolk Horse Society n.d.). In contrast, the breeding goal of the Thoroughbred (TB) racing industry is to breed horses which are capable of achieving top race speeds and placings, therefore breeding selection programs are based on the presence of these phenotypic traits which are available from performance records. Race speed and earnings traits have been successfully used in TB breeding programs to inform horse selection, increasing the likelihood of genetic progress (Aceto and Perente 2012, Thiruvenkadan et al. 2009 and Buxadera et al. 2008). In the sport horse breeding industry, the generic breeding goal most commonly emphasised in literature is to produce competitive horses at the advanced levels of competition (Hellsten et al. 2006).

The term sport horse is a collective term used to describe a horse bred for either showjumping (SJ), dressage, or eventing (a combination of the two and cross country) (Stewart et al. 2012 and Hellsten et al. 2006). Each sporting discipline requires different physical demands from the horse, therefore the phenotypic traits assessed in sport horse evaluations vary depending on the discipline the horse was bred for (Hellsten et al. 2006). Stewart et al. (2010) maintains that key traits of the dressage horse are a suitable temperament and athleticism to exhibit

ride-ability. Desired traits of the showjumping horse are a powerful jump and ride-ability (Viklund et al. 2010a). The eventing horse is said to have the versatility to perform with relaxation and suppleness in the dressage phase; with speed, strength, stamina and athleticism in the cross country phase, and lastly with co-ordination in the showjumping phase (Back and Clayton 2013 and Stewart et al. 2012). Koenen et al. (2004) determined that the phenotype traits that are most commonly prioritised and assessed in different studbook breeding programs are horse conformation and locomotion and that behavioural and fertility traits are the least emphasised.

Horse conformation and locomotion have been linked to horse performance ability for decades. In 1993, Holmstrom et al. studied quantitative conformational measurements of 195 performance horse to assess what body measures were most desirable and concluded that characteristics such as a long sloping femur, inclined scapular and short cannons were highly correlated to horse performance ability. More recent research by Back and Clayton (2013) linked high dressage competition scores with horses with small hip angles and long sloping femurs. In showjumping, high competition scores have been linked with the quality of the canter (Back and Clayton 2013 and Hellsten et al. 2006). Back and Clayton (2013) suggest that the ideal stride frequency of a jumping horse is between 108 and 157 strides per minute, with a stride length of approximately 4m, shortening down to 2.39m before take-off for competition success. Limited research is available which describes the ideal phenotypic traits of Eventing horses, however it can be suggested that a combination of dressage and showjumping traits are required. In addition, during the cross country phase of competition, the eventing horse requires a strong galloping stride and jump ability to travel a distance of up to 6270m at an average speed of 9.5m/s and jumping obstacles (Munsters et al. 2014 and British Eventing 2012). Whilst research clearly demonstrates that horse conformation and locomotion correlate with performance ability, disagreements are present in literature regarding which specific traits are the most advantageous across disciplines. In 1993, Holmstrom et al. linked horse height with performance. More recent research by Stewart et al. (2010) was in agreement with Holmstrom's et al. (1993) findings and suggested that horses which measured 180cm in height were the best performers. On the other hand, Jonsson et al. (2014) state that horse height of 163-169cm is optimal for performance as this range positively correlates to career longevity due to a relatively larger cannon bone circumference associated with a smaller horse. Both Jonsson et al. (2014) and Bowing and Ruvinsky (2000) have associated cannon bone conformation with horse durability and career longevity. Therefore in summary, current research suggests that conformation and locomotion traits constitute important criteria when selecting horses for optimal performance ability, however these traits

must be selected for both durability and performance to increase career longevity and ultimately success.

Whilst the concepts of desirable conformation and locomotion traits of the sports horse are apparent from published research, Koenan et al. (2004) state that the subjective and generalised terms used to describe these phenotypic traits by studbooks limits the achievement of consistency between sport horse breeding objectives and slows genetic progress of the breed (see Appendix 1a for the detailed breeding goals of 19 studbooks). However unlike many, the breed specification given in the Norwegian Warmblood studbook is more objective, stating that their Warmbloods must be between 1.60 and 1.70m in height and must be selected on 40% conformation and 60% performance capability, which includes soundness and temperament, although some subjective terminology is still used (Koenan et al. 2004). Clear, consistent and objective breed specifications are needed to optimise breeding selection across the many operating studbooks; to make this possible more concise and specific definitions of the desired biological traits are needed (Koenan et al. 2004). It is widely agreed that whilst the breeding objectives of sports horses vary by studbook organisations, the common goal is to breed superior competition horses who can perform successfully at the advanced levels of competition. Therefore phenotypic evaluation programs seek to select and breed superior horses with desired highly heritable traits to increase accuracy in estimations of merit (Hellsten et al. 2006). The information used in sports horse breeding to identify superior horses is formed from pedigree details combined with performance data (Hellsten et al. 2006), similar to in the Livestock and Working dog industries. Pedigree data alone have been used to demonstrate equine diversity, but not to evaluate performance ability (Roos et al. 2015 and Hamann and Distl 2007). The performance data which are used in phenotypic evaluation of the sport horse are collected from a variety of sources such as adult horse competition results, young horse competition results, station tests and 1 day field performance tests.

Phenotypic evaluations using records from competition results

Competition results are used to form phenotypic evaluations of the horse and are collected from competition placings, scores, penalty points and monetary winnings. Braam et al. (2011) investigated 17,962 horse competition results to evaluate dressage, showjumping and eventing performance ability of horses which were at least 12 years old. The variable of horse age is important to control for as it is highly associated with performance ability (Hellsten et al. 2006). The results of Braam et al.'s (2011) study suggest that horses with the most competition placings had the longest career (heritability 0.17, SE 0.02). However, competition placings in Sweden are only awarded to the top 20% introducing selection bias and skewed results, which

demonstrates the importance of an unbiased and reliable performance measure. A further performance measure is horse competition penalty points; these are awarded to all competing horses, reducing selection bias. Horse penalty points are awarded in eventing competitions and can therefore be used to evaluate the heritability accuracy of each phase of competition (dressage, showjumping and cross country). Kearsley et al. (2008a) and Stewart et al. (2012) investigated the use of competition penalty points of eventing horses to achieve performance trait heritabilities which represented a single discipline. Results demonstrated that the highest heritability estimates were achievable from the showjumping phase (0.15 and 0.22), then the dressage phase (0.10 and 0.17) and the lowest from the cross country (0.03). These ranges of heritabilities estimates across the phases of competition demonstrate the effect of discipline on performance and the accuracy of the genetic information available from heritability estimates. Stewart et al. (2012) further demonstrated that by analysing the higher levels of competition only (advanced), heritability estimations of performance increased; improving test reliability. However if evaluations were based on advanced competition results only, the generation interval would be too large as horses would need to mature before their records could be evaluated. It is apparent that whilst positive heritability estimates are achievable from performance measures, they can be low in accuracy and as such limit the genetic information available for breeding success. Additionally they can take a long time to obtain, increasing the generation interval and consequently slowing the genetic progress of the breed. In horse breeding the shorter the generation interval is, the quicker breed progress is as superior horses can be identified early on in their careers for training or breeding purposes.

In France, the breeding program of the Selle Francais is informed by specific young horse competitions, entitled Circuit Classique, to reduce the generation interval. Dubois and Ricard (2007) reviewed the Circuit Classique and summarised that positive performance heritabilities of 0.14 were predictable from annual earnings of horses aged 4 to 6 years. Whilst these heritability estimates are low in available phenotypic information, they do suggest that the generation interval can be decreased by analysing horses aged 4 to 6. However more recent research by Posta et al. (2010) and Stewart et al. (2012) suggest that since age is significantly associated with performance ($P<0.001$) and performance peaks at approximately 10 years old, competition results of the young horse may not be completely representative of their ability. Posta et al. (2010) demonstrated that for horses 10+ years old, performance heritabilities of 0.38 were possible, which is considerably higher in estimates of phenotypic information than from the Circuit Classique; demonstrating the importance of horse age. However it is important to note that the discrepancy between Posta et al. (2010) and Dubois and Ricard (2007) findings may have resulted from the different competition measures used (winnings/faults) to evaluated performance in their research. Whilst phenotypic evaluations of

the young competing horse are limited, France and Belgium rely predominantly on these measures to inform their breeding practices and as such 45% of their Selle Francais horse population have been evaluated, increasing the pool of data available for analysis (Hellsten et al. 2006). A large pool of available data can assist selection intensity, increasing genetic gain. The term selection intensity refers to the standardised measure of the superiority of the parents in comparison to the mean of the population (Falconer and Mackay 1995). Therefore the greater number of horses tested within a population increases the pool of phenotypic information available, consequently resulting in a greater number of animals to select from, therefore leading the superior animals to be selected for breeding and increasing the odds of achieving genetic gain. It is clear that whilst using competition results to form performance evaluations of the young horse allows for greater pool of phenotypic data to become available, there are some limitations in the form of low heritabilities. Additionally evaluations based on competition results only measure performance in one discipline which the horse has been trained for, introducing selection bias, but more importantly these phenotypic evaluations slow genetic progress by increasing the generation interval.

Phenotypic evaluations using records from station tests

Germany and The Netherlands rely on young horse performance tests to collect phenotypic information; they have a much shorter generation interval in comparison to France and Belgium and additionally are highly placed in the international studbook rankings (WBFSH 2015a, WBFSH 2015b, WBFSH 2015c). Young horse performance tests can be split into two categories, station tests and field performance tests. Station tests generally consist of uniform testing of certain phenotypic traits which are thought to predict the potential ability of a young horse, such as conformation and locomotion. Horses are entered into discipline/purpose specific tests and scored subjectively by judges on their suitability for purpose. Station tests usually last about 1 to 3 months, whereas field tests usually last 1 day and are one-time judgments of the horse (Hellsten et al. 2006). The specific length of the test, age of horse and scoring criteria varies depending on the judging organisation (Koenen et al. 2004), however Hellsten et al. (2006) reviewed extensive research of European station testing and states that whilst tests vary, performance phenotypic heritabilities on average were found to be moderate to high in information available (Luehrs-Behnke et al. 2002b, Gelinder et al. 2001 and Olsson et al. 2000).

Station testing in Germany is used to inform training and breeding decisions of the sport horse. Luehrs-Behnk and Kalm (2002a) conducted an extensive review of Germany's stallion performance test (SPT) and mare performance test (MPT) with 4,527 and 40,670 horse records of horses aged 3 to 4. The SPT lasts 70 days, whereas the MPT lasts 21 days,

however similar traits are assessed such as the horse's gaits, ride-ability and free jumping ability. Luehrs-Behnk and Kalm (2002a) concluded that heritabilities for the SPT ranged from 0.33 to 0.51 (SE 0.02) and 0.27 to 0.38 (SE0.01) for the MPT; test duration may contribute towards the lower performance heritability estimates of the MPT as the longer the test the more accurate phenotypic information can be derived to increase estimations. However, research by Olsson et al. (2000) and Gelinder et al. (2001) have estimated heritabilities ranging from 0.32 (SE 0.1) to 0.55 (SE 0.08) for 7 day station tests, suggesting that the length of test may not entirely influence estimation accuracy and instead variables such as test criteria, horse age and the number of assessed horses may be influential (Hellsten et al. 2006). Number of horses evaluated is important for phenotypic evaluation reliability as it increases data availability. On average, European station testing attendance is between 10% and 20% of registered foals (compared to 45% in France and Belgium); the cause of this can be contributed to the labours and costly test design (Hellsten et al. 2006). Therefore it can be concluded that whilst station tests shorten the generation interval and predict higher biological trait heritabilities than competition results, attendance is considerably less. Gelinder et al. (2002) suggest that the more accessible a performance test is, the more likely people will attend, resulting in greater data availability. Therefore the practicality of 1 day field performance tests (FPT) over station tests can be beneficial to sport horse breeding.

Phenotypic evaluations using records from field performance tests

FPT are another method used in the phenotypic evaluations of young sport horses. FPT take place in one day and depending on the organisation running the test, these tests usually take place in varying locations, making the test accessible to participants; encouraging attendance and participation. Hellsten et al. (2006) demonstrated that population attendance of European FPT ranged from 10% to 43%, which is greater than that of station testing and therefore increasing data availability and consequently test reliability. FPT resemble station testing as the phenotypic traits assessed at evaluations are those which are thought to be predictions of performance ability (conformation and locomotion). Similarly, horses are entered into discipline/purpose specific tests and scored subjectively by judges on their suitability for purpose. However, it can be suggested that as a result of the less comprehensive testing design and environmental influence of varying locations, predicted biological trait heritabilities may be lower than station testing estimations. Hellsten et al. (2006) demonstrated performance heritabilities of station testing of 0.42, in contrast heritabilities of 0.32 were demonstrated for FPT which suggest that the observations result in less phenotypic information and as such will deliver less accuracy in estimations of merit. In agreement with Hellstens et al.*'s* (2006) results, Viklund et al. (2008) demonstrated that Swedish FPT performance heritabilities averaged 0.3.

The effect of test design on heritability estimates can be demonstrated by reviewing the work of Viklund et al. (2008) who estimated FPT performance heritabilities and concluded that the fixed effect of event location had the greatest significance on test results compared to the effect of horse sex or age. Event location incorporates the effect of the judge, weather and environment (for example indoor versus outdoor testing location). The range of judges used can influence horse scores as they are awarded subjectively and therefore may be influenced by personal opinion (Hellsten et al. 2006). Furthermore, as FPT take place on different days in varying locations, evaluated horses will be subject to different weather conditions and as such the behaviour and performance of the horse may be affected, influencing their scores. Jorgensen and Boe (2007) demonstrated that moderate to heavy rain and wind conditions correlated with increased restlessness displayed by horses. More recently, Iwona et al. (2015) measured the heart rate and body temperature of the horse and concluded that increased wind speeds (<5.5m/s) correlated with an increase in horse physiological responses and adverse behaviours, such as unwillingness to work. Therefore as FPT assess the performance of the horse on the day, it can be suggested that the test day conditions may influence the scores awarded. Comparably, station tests take place in one location, over greater time periods with consistent judges, demonstrating that FPT are subject to more test variables (Hellsten et al. 2006). To increase test integrity, Suontama et al. (2009) suggest decreasing testing subjectivity by introducing quantitative testing measures.

Suontama et al.'s (2009) research can be examined to establish the differences in heritability estimates of subjectively and objectively scored horse conformation to establish which method can deliver the richer estimations of phenotypic information to help deliver greater accuracy in estimations of merit, which were 0.15 and 0.65 respectively (average horse age 6.6years). Additionally Suontama et al.'s (2009) and Schroderus et al.'s (2006) research can be compared to establish the differences in heritability estimates for different horse birth groups to establish which method can deliver the richer estimations, which were 0.15 for horses aged 7 and 0.25 for 1 to 3 year olds. The limited information available from subjectively judged conformation of 7 year old horses may result from environment factors such as nutrition and training which can influence their appearance more than that of 1 to 3 year olds and as such influence the accuracy of information available for phenotypic evaluations. In many FPT's, conformation is judged by visual inspection alone and has demonstrated heritabilities estimations of 0.28/0.23 for 3 and 4 year olds respectively (Viklund et al. 2008) and 0.25 for horses aged 1 to 2 (Olsson et al. 2008), demonstrating that the age and condition of the horse can influence the availability of phenotypic information, influencing the accuracy of breeding on merit. Quantitative measures of conformation are currently used in Icelandic FPT and have demonstrated increased availabilities of phenotypic information (heritability 0.43) for horses

aged 4 to 6 years old. Therefore, although it is clear that quantitative assessed conformation can assist in increasing the information available for breeding on phenotypic merit, there are some limitations associated with this method. In Albertsdottir et al.'s (2008) study, the conformational characteristics of a single population of Icelandic horses which have a uniform conformation breeding goal were assessed, whereas in sport horse populations the goals are varied and therefore this implementation of system would be more complex. For instance, Berrey et al. (2002) evaluated the conformation characteristics of dressage horses and concluded that conformation variables differed between breeds, with German bred horses being more adapted for dressage. Hellsten et al. (2006) suggests that in order to limit FPT subjectivity, judging must be standardised by clearly defining the subjective scoring criteria.

Test subjectivity can be further reduced by minimising the variables present in the test design. Viklund et al. (2008) demonstrated the effect of a rider/handler by comparing the results of individual FPT aimed at 3 and 4 year old. These FPT's were similar in design, however the 3 year old test includes no ridden elements, whereas 7 ridden element are in the 4 year old test. Viklund et al. (2008) concluded that trait heritability estimations averaged 0.36/0.33 for 3 and 4 year olds respectively, suggesting that by examining the horse alone, the analysis is higher in phenotypic information, increasing greater accuracy in estimations of merit. Horse age has been demonstrated to increase phenotypic information estimations (Posta et al. 2009 and Stewarts et al. 2010), therefore a further suggestion for the lower estimations of 4 year olds in Viklund et al. (2008) study is the effect of the increased presence of a handler/rider on the horse's performance. The effect of the rider variable can be further demonstrated when comparing various FPT literature. Greater information is available from phenotypic FPT's that assess the trait of free jumping, in comparison to limited information from assessing the trait with a rider, suggesting that by removing the rider variable greater accuracy in estimations of merit are possible (Becker et al. 2011, Olsson et al. 2008, Viklund et al. 2008 and Ducro et al. 2007). An explanation for this may result from the understanding that rider experience can influence horse performance and impression in testing conditions; consequently influencing the subjective scores awarded by judges (Hellsten et al. 2006). Research suggests that to increase test integrity, the rider variable can be taken out of the FPT design, as scores from 3 year old FPT (which include no ridden elements) were highly correlated (0.82 to 0.99) to those of 4 year olds (Viklund et al. 2008). Ultimately however, the chosen traits tested in sport horse evaluations must correlate highly with later competition results as the overall breeding goal is to produce competitive horses (Hellsten et al. 2006).

Phenotypic evaluations using combined records from field performance tests and competition results

In the Livestock industry, the accuracy of phenotypic evaluations has been increased by using more than one source of phenotype information (Tsuruta et al. 2011, Sun et al. 2010 and Negussie et al. 2006). In the Equine industry, many European countries have correlated field performance scores with later competition results to determine test reliability. Dutch First Stallion Inspection tests (FSI) take place in one location and day, with the same judges and have been shown to moderately correlate scored gait scores to dressage competitions point results (0.55) and highly correlate jumping scores to showjumping competition point results (0.80), which may be a result of the more objective scoring in showjumping competitions (Ducro et al. 2002). Interestingly, Ducro et al. (2007) demonstrated lower correlations of 0.52/0.72 for dressage and showjumping competition points and traits scored at the Dutch Stud Book Entry inspection tests which take place in one day at different locations with different judges, which makes evident again that test accuracy is slightly reduced with the added variable of test location and judges. The Swedish Riding Horse Quality Test (RHQT) is a one day young horse performance test which attracts the highest population of sports horses compared to other European FPT's at 43% (Hellsten et al. 2006), increasing the pool of phenotypic data available for analysis. Substantial research has been undertaken on the RHQT to demonstrate its effectiveness in the Swedish sports horse breeding programme.

Wallin et al. (2003), and more recently Viklund et al. (2010a) established the RHQT effectiveness by correlating the test trait scores of gaits and jump ability to competition results (cumulative points and placing) in dressage and showjumping. Wallin et al. (2003) and Viklund et al. (2010a) demonstrated correlations of 0.69/0.62 for gait scores and dressage performance and 0.88/0.84 for jumping scores showjumping performance, suggesting that the RHQT is successful in assessing young potential dressage and showjumping horses and as such can lead to genetic progress. No published peer reviewed research is available on correlations of the RHQT and eventing competition results, however in a Master's dissertation by Ray (2012), RHQT scores were correlated with eventing results (lifetime points and placing), demonstrating that correlations of 0.17/0.35 were possible. Ray (2012) additionally demonstrated that eventing competition results moderately correlated with showjumping competition results, which is logical as 2 phases of eventing have a jumping element and as 47% of Swedish eventing horses also competed in showjumping competitions, phenotypic improvements to showjumping horses should improve the eventing horse.

Olsson et al. (2008) broadened Swedish research by correlating the Swedish stallion performance testing (SPT) with competition results and RHQT data. Correlations between

SPT gait and jump scores with competition results in dressage and Showjumping were 0.60 and 0.87 respectively, suggesting that the SPT is very reliable in its evaluations. In all Swedish studies, correlations between jumping traits and showjumping performance are the highest demonstrated which can result from the more objective scoring of jumping ability in performance testing and in competitions, compared to the subjective scoring of horse locomotion and dressage performance (Hawson et al. 2010). Olsson et al. (2008) furthered in her research that correlation accuracies could be increased by 34% for dressage and 13% for showjumping by incorporating RHQT and SPT data with competition results, which indicates that by combining all available measures of performance testing, breeding selection can take place with more reliability; increasing genetic progress. Furthermore Viklund et al. (2010b) has demonstrated that richer heritability estimations and lower residual variances were obtainable from more recent RHQT evaluations (1988 to 2007 compared to 1973 to 1986). As residual variances have decreased from 1973 to 2007, this suggests that some genetic progress of the sport horse breed has been made as breed characteristics have become more consistent to breed specifications. In addition these results suggest that the judging of RHQT horses has improved over time. It is clear from research by Wallin et al. (2003), Olsson et al. (2008) and Viklund et al. (2010a) that the RHQT results are good indicators of later performance at competition and the RHQT can be used to facilitate the genetic progress of Swedish Sports horses. Furthermore, in Sweden 75% of the sports horse population compete in showjumping, 40% in dressage and 10% in eventing, therefore 20% compete in more than one discipline (Olsson et al. 2008). Showjumping results have demonstrated the highest correlations with RHQT scores, therefore it is apparent that the reliability of breeding selection of showjumping horses is greater, which supports the Swedish industry trend of producing more competitive showjumping horses. In conclusion, whilst European countries may use different tests to support sport horse breeding programs, they all use phenotypic evaluations as a fundamental tool in their industries.

Sport horse breeding in Europe

Phenotypic evaluations of the young horse can be used to predict potential ability in varying disciplines, reducing the generation interval and supporting genetic gain (Hellsten et al. 2006). The World Breeding Federation for Sport Horses (WBFSH) calculate studbook rankings based on the competition points earnt by registered and competing horses in different countries. The WBFSH studbook rankings make apparent the success of the breeding programmes employed in Germany, the Netherlands and Sweden (see Table 1). Germany and Netherlands repeatedly have repeatedly dominated competition rankings, with Great Britain (GB) performing to their best in eventing competitions, but losing out to European countries in dressage and showjumping (see Table 1).

Table 1. WBFSH Studbook rankings (WBFSH 2009 and WBFSH 2005).

WBFSH studbook rankings

Year	Dressage		Showjumping		Eventing	
	2015	**2004**	**2015**	**2004**	**2015**	**2004**
Great Britain	19th	Not placed	34nd	21st	6th	6th
Germany	2nd	1st	1st	3rd	2nd	3rd
Netherlands	1st	2nd	2nd	1st	5th	11th
Sweden	8th	6th	10th	9th	12th	9th

Counties can have more than one competing studbook. In this table, only the highest placing studbook is used.

Crossman (2010) investigated the structural framework of various European equine industries and concluded that German and Swedish industries are the most organised in design and subject to high levels of government involvement, resulting in strict regulations on breeding stock. Barrey et al. (2002) supports this by demonstrating that German breeds are most adapt for competition performance in dressage. Furthermore, in Sweden, the leading studbook the Swedish Warmblood Association represents the breeding of all sports horses and endorses standardised phenotypic evaluation tests (Crossman 2010). The RHQT is an effective tool in Swedish sport horse evaluations (Viklund et al. 2010a, Olsson et al. 2008 and Wallin et al. 2003). In contrast, Crossman (2010) describes Great Britain's (GB) structure as fragmented with low levels of government involvement (see Appendix 1b). Many studbooks in GB operate independently of each other and this lack of standardisation can weaken breeding programmes (Koenen et al. 2004), which is evident from the studbook rankings in Table 1. The Sport Horse of Great Britain (SHB) and the British Warmblood studbooks aim to improve and promote British bred horses to achieve higher recognition on the WBFSH rankings (the British Warmblood 2015 and SHB 2013). Therefore developments are necessary in the British breeding industry in order to achieve this genetic progress of the British sport horse.

Sport horse breeding in Great Britain

In Great Britain, the breeding aim for the sport horse is to produce athletic horses which have the potential to succeed at showjumping, dressage and eventing (SHB 2013 and The British Warmblood 2015). Additionally the breeding aim is to produce horses for amateur riders, which are the majority of the riding and competing population; the British Equestrian Trade Association's (2015) national equestrian survey highlighted that 96% of UK equestrians rode for pleasure and 59% of these competed in unaffiliated competitions. Although the sample size of this survey is unclear and therefore the results must be taken as estimates. When breeding horses for amateur riders, phenotypic traits such as temperament and general all round ability are highly important (Stewart et al. 2012 and Kearsley 2008b). When breeding horses for professional competition, phenotypic evaluations based on horse competition

performance can be utilised, as used in France, Belgium and Ireland (Hellsten et al. 2006). As the accuracy of phenotypic evaluations is dependent upon the number of records used (Stewart et al. 2012), it is important that all available records are incorporated into phenotype analysis. The governing bodies of UK sport, British Eventing, British Showjumping and British Dressage all hold comprehensive horse competition results databases for their respective disciplines and therefore these data can be used to inform breeding practices of British sport horses. Research into the use of phenotypic evaluations in the UK is clearly developing.

Stewart et al. (2012) and Kearsley et al. (2008a) investigated the competition data of British Eventing horses to estimate performance heritabilities using penalty point scores from each phase/grade of competition. Previously eventing heritabilities of 0.11/0.17 (SE 0.01) have been estimated using annual results (horse starts, places and rank), suggesting that limited phonotypic information is available to achieve high accuracy to breed on merit (Ricard and Chanu 2001). However, by analysing each competition phase, the heritabilities estimations were higher for dressage (h^2 0.24 and SE 0.02), showjumping (h^2 0.31 and SE 0.01) and for cross country (h^2 0.52 and SE 0.008) (Stewart et al. 2012). In comparison, by altering the analysis design and including the rider variable, lower estimations were predicted of 0.10 for dressage, 0.15 for showjumping and 0.03 for cross country (Kearsley 2008b), suggesting that these observations are lower in phenotypic information. Including the rider variable in estimations of horse performance can be beneficial as it this relation between the horse, rider and competition performance that needs to be understood to provide information to increase greater accuracies of selecting and breeding successful sports horses. Furthermore the studies Kearsley et al. (2008a) and Stewart et al. (2012) both demonstrated that young horse results at the lower grades of competition correlated strongly to the higher grades, suggesting that young horse results can be used to predict future ability. Therefore whilst eventing competition heritabilities provide limited phenotypic information, they are all significantly different from zero and can potentially be used in British horse selection programs. Furthermore Kearsley et al. (2008a) and Whitaker et al. (2004) investigated the effects of competition variables and demonstrated that the event location had a significant influence on eventing horse scores and variance, which is understandable as the terrain and technicality of courses vary between facilities. However it must be noted that course designs meet the standardised regulations set by British Eventing (British Eventing 2009). Additionally Kearsley et al. (2008a) demonstrated that the effect of rider and horse age effect was greater as the grades progressed. Horse sex had no effect (Kearsley et al. 2008a and Whitaker et al. 2008). As the influence of competition variables increases with the increasing grades of competition, it can be suggested that accurate selection must be made of the horse at the lower levels of competition to account for the influence of fixed effects at the higher levels.

Stewart et al. (2012) suggest that British breeders are placing more selection emphasis on cross country performance, over dressage and showjumping of eventing horses, instead research has suggested that greater emphasis should be placed on dressage and showjumping ability. Research by Kearsley et al. (2008a) and Whitaker et al. (2004) demonstrated that penalties accrued in the dressage phase of competition accounted for 60/80% and 63% respectively of overall penalties, resulting in strong correlations between dressage penalties and final competition placing. However, Kearsley et al. (2008a) furthered that dressage scores only correlated by 0.13/0.33 (SE 0.06/SE 0.08) with the showjumping and cross country phases of competition, suggesting that eventing horses should not be selected only on dressage ability, additional phenotypes are required for the other phases. Stewart et al. (2012) research demonstrates that correlations between the grades (levels) of eventing competition were 0.59/0.99 for dressage and 0.74/0.99 for showjumping, implying the same phenotypes are required for low and high level dressage and showjumping performance. Therefore when selecting horses for eventing, emphasis should be placed on selecting good dressage and showjumping horses as these are more predictive of performance at the higher grades.

The British Showjumping database was examined in Stewart's (2012) PhD thesis to understand its parameters. From competition results of placing and penalties, heritability estimates were 0.05/0.08, which suggests less phenotypic information is available from this discipline when compared to the showjumping phase of eventing competition (0.31) (Stewart et al. (2012). Further to eventing and showjumping, the parameters of British dressage horses have been estimated. Stewart et al. (2010) used the percentage mark scores awarded to horses at dressage competitions and demonstrated that some phenotypic information can be obtained from these scores (0.11/0.15 and SE 0.02), which additionally is in the same range as Kearsley et al. (2008a) predicted for the dressage phase of eventing competition. Additionally, Stewart et al. (2010) research suggests horse breed significantly influences performance, with warmblood horses performing the best, which is not surprising given the intended goal of the breed. Age was also significantly associated with performance, suggesting that young dressage horses can be selected on their competition results to predict potential performance, reducing the generation interval and increasing progress of British breeds. Furthermore, Stewart et al. (2010) demonstrated that some genetic progress of British sport horses has already taken place as EBV's (calculated by BLUP) have increased in value and reliability in the past 15+ years (1985 to 2001). The WBFSH (2015a) studbook rankings support this trend as in 2004 British studbooks were not placed in the dressage rankings, however in 2015 they were 19th out of the 40 which were placed. Whilst competition data can inform phenotypic evaluations and better breeding practices of British sport horses, the

extensive generation interval is a limiting factor as the horse must reach the age of 4 to begin competing. Horses are sexually mature at 12 to 15 months, therefore it is possible to breed them before phenotypic information is available from competition results, reducing the accuracy of selection (EquiMed 2010).

Field performance testing of British sport horses

Genetic progress of British sport horses is supported by the British Equestrian Federation (BEF), the National Governing body of horse sports in the United Kingdom (BEF 2015a). The BEF in 2003 implemented the British Breeding initiative in an attempt to co-ordinate and develop the sport horse studbooks that are operating in Great Britain (Crossman 2010). The initiative was the introduction of a young horse field performance test, titled the Young Horse Evaluation series (YHE). The YHE took place in Britain in the summers of 2002 to 2005; tests were completed in one day and in varying locations. British 4 to 6 year old sports horses which entered YHE were judged by a panel of three judges on the suitability of their performance traits for their chosen discipline (dressage, showjumping or eventing). These performance traits are conformation, paces, loose jumping ability, and ridden jumping ability. Additionally, a veterinary examiner would assess the horse's general health and soundness. The scores, which were subjectively awarded by judges, were averaged and used to represent the horse's potential ability (these scores being from 0 to 10). The YHE was comparable to the Swedish RHQT, as both tests were of a similar design (Stewart 2012). The RHQT has been proven beneficial to the sports horse breeding industry of Sweden (Viklund et al. 2010a), however the same has not been demonstrated for British YHE (Kearsley 2008b).

Kearsley (2008b) investigated the parameters of the YHE by measuring the complete dataset of 248 individual horse scores. Gender had a significant effect on horse scores with stallions receiving the highest scores; this can be a result of greater selection criteria being placed on stallions used for breeding. Mean scores differed significantly across the dates of evaluation, which suggests that the effect of location and judging team greatly increased the variance of horse scores. It was found that the YHE's was significant in assessing horse variation ($P>0.05$), with moderate correlations being present between the assessed traits. However heritabilities estimated in Kearsleys (2008b) were not significantly different from zero and therefore cannot be used to inform horse breeding programs as very limit phenotypic information is available, which questions the usefulness of the YHE. The non-significant heritabilities can be linked to a lack of sufficient data, only 248 horse records were available, in comparison to the 3,708 records used to estimate the parameters of the RHQT (Wallin et al. 2003). British breeding recognised the restrictions of the YHE; it was found that participation was limited as a result of the induction of young horse classes for 4 year olds by competition

bodies (British Breeders Magazine 2009). In 2005, British Breeding developed the YHE into the Futurity scheme, which is open to foals, yearlings, 2 and 3 year olds, which increased participation and accessibility as anticipated, whilst reducing the generation interval (Kearsley 2008b).

The Futurity scheme aims to identify talented young British bred horses to inform sport horse breeding and training in the UK by providing data for phenotypic evaluations. The Futurity test is for British bred horses only and is open to foals, yearlings, 2 and 3 year olds (see Appendix 1c for further details on eligibility BEF 2015b). Horse age is a test variable which can influence the information derivable from FPTs and research has demonstrated that there are benefits of assessing each individual age group. Tavernier (1992) demonstrated that trait phenotypic information can increase with age (from 0.67 for 4 year olds to 0.85 for 6 year olds). On the other hand, Schroderus *et al.* (2006) and Suontama et al. (2009) demonstrated that judging subjectivity can be reduced and accuracy increased by testing 3 year old horses as they are less influenced by environmental factors such as nutrition and training. Viklund et al.'s (2008) agrees with testing 3 year old horses as their scores highly correlate with the scores of 4 year olds (0.82 to 0.99) and the generation interval is decreased. Therefore it is evident from research that the inclusion of 3 year olds and younger in the Futurity test may have a positive result on test usefulness as the generation interval is decreased and attendance can increase, however this is yet to be investigated.

During Futurity evaluations, horses are presented by their handler (not ridden) and scored subjectively on discipline specific traits on a scale of 1-10 by 3 Futurity judges and a veterinary examiner. The test disciplines are dressage, showjumping, eventing, endurance and sports pony. The overall score awarded to evaluated horses represents a premium, the higher the premium, the higher the competition level the horse is said to be suitable for (see Appendix 1d). Horse premiums have become increasingly more detailed over time (BEF2015c and BEF 2014a) and for a comparison of the scoring criteria, see Appendix 1d and 1e. Furthermore trait criteria assessed in the Futurity has also undergone developments; in 2006 and 2007, the trait "type" was assessed, this criteria changed to "type and temperament" in 2008, but was then removed as it could not be easily and objectively assessed (Kearsley 2008b). Traits scored in 2014 were; conformation, correctness of paces for discipline, athleticism and jump ability (for horses aged 3 and being assessed for eventing or showjumping) (BEF 2014b see Appendix 1f for an example of trait scoring criteria). The veterinary mark which evaluated the health and soundness of the horse was taken out for the 2006 judging season but reintroduced in 2007. Whilst the Futurity have clearly undergone some changes, the phenotypic traits currently scored in the Futurity have all demonstrated usefulness in other European FPTs. Furthermore these traits have correlated positively with competition performance (Viklund et al. 2010b,

Ducro et al. 2007 and Wallin et al. 2003), which justifies their inclusion in the Futurity scheme. However in European studies, the trait score of orthopaedic health, which is similar to the Veterinary score used in the Futurity, has demonstrated low heritabilities (Jonsson 2013, Viklund et al. 2008 and Stock and Distl 2006) and negative correlations to competition results (Wallin et al. 2003). This is an interesting result as it suggests the higher (better) horses score at FPT's the worse they perform in competition. This negative correlation implies that as sport horse breeding goals primarily focus on performance success (Koenen et al. 2004) welfare issues may result and as such limit the horse's health and longevity. Research by Velie et al. 2015 and Braam (2011) suggest that durability traits of the horse can equal heritability estimates of other performance measures and as such can be implemented into selecting and breeding programmes to improve the welfare of the horse.

Temperament is a trait which is no longer assessed in Futurity evaluations as it was thought that it was an unfair and unreliable assessment of the behaviour of the young horse due to the tests taking place in an unusual environment (Dixon 2015 personal communication). However a recent review of horse temperament and performance by Randle et al. (2015) suggests that horse temperament is an important trait of the competing horse as behaviour can influence performance. Furthermore in dressage tests, marks are awarded for the submission displayed from the horse, highlighting the importance of temperament during competition (Randle et al. 2015). Although low heritabilities information of 0.17 (SE 0.24) have been demonstrated by Rothmann et al. (2014) for the trait of behavioural reactivity during one day FPT's and heritabilities of 0.08 (SE 0.04) during ridden performance tests suggesting first that limited phenotypic information for breeding is achievable from the trait. Research by Visser et al. (2003) demonstrated that there is no desired single behavioural trait of Showjumping horses as individual riders value different behaviour traits of their horses. Therefore as a result of the complex nature of measuring and testing horse temperament, and the complexity of matching horse temperament with rider preferences, the inclusion of a trait for temperament in the Futurity would be multifaceted in definition which could be suggested reduce the value of evaluations, however this has not been demonstrated.

In 2015, the Futurity underwent further changes to the assessment design and implementation to increase test objectivity and standardise horse scores. Linear scoring has been introduced which was adapted from systems used in Germany (Stock 2013) and the Netherlands (Viklund 2010a). Linear scoring is considered to make scoring more objective as horses are scored compared to the norm of the population with the use of more descriptive criteria (BEF 2015d). Test objectivity is also said to be improved by changing the evaluator panel; judges with more international competitive experiences have been introduced, although the benefits of this have yet to be demonstrated. The Futurity has clearly undergone some significant developments

overtime, leading to some inconsistencies in recorded data, therefore limiting its transparency. However, whilst changes have taken place, the underlying process ultimately remains unchanged and similar to FPTs used on the continent; young horses are evaluated on their phenotypic suitability for purpose and awarded a corresponding evaluation score. European young horse tests have been proven effective (Hellsten et al. 2006), however due to the subjective nature of FPTs and the variables that exist (including location, horse age, trait scoring criteria and horse temperament), it cannot be presumed that Futurity evaluations are as effective and it is widely agreed that this needs to be investigated (Stewart 2011, Horse Breeders Magazine 2009 and Kearsley 2008b).

The BEF Futurity: Current Research

Kearsley (2008b) first attempted to investigate the parameters of the YHE and Futurity test. From YHE data, n=248 horses had evaluation records and from the Futurity data, n=72 horses had evaluation records (these Futurity records were from evaluations in 2005 only). Results from the YHE dataset and the Futurity dataset demonstrated moderate correlations between the assessed traits (0.69 and SE 0.08), however due to the small dataset no more significant results were demonstrated. Stewart's (2011) PHD thesis further investigated the parameters of the Futurity scheme, with the use of data from 2006 to 2009, equating to 1887 records of evaluations, which took place in 15 locations over 35 evaluation days. There were some inconsistency with the data, for example, no veterinary trait was scored in 2006, in 2008 the trait "type" became "type and temperament", in 2009 the trait "correctness of pace" was introduced and the definitions of the assessed traits have become more precise and detailed over time. Stewart's (2011) results suggested that heritabilities ranged from 0.2 to 0.42 (SE 0.2) for the assessed traits, with athleticism achieving the lowest and type and temperament achieving the highest heritabilities, which questions its removal from the judging criteria? Furthermore, in the 2014 scoring criteria, the trait of athleticism had greater weighting on the overall score as the score from this trait is multiplied by 2, whilst all other traits were scored out of 10, therefore this reasoning is questionable as it is the trait with the lowest heritability estimation in Stewart's (2011) study. The heritability for the trait of conformation equalled 0.29, which is in a similar range to the RHQT of 0.33 (Wallin et al. 2003). This is a promising result for the Futurity evaluations as it suggests that the Futurity is operating similarly to the effective RHQT (Viklund et al. 2010b). Correlations between all traits (apart from veterinary score) were high, indicating the same genes are largely responsible for all the phenotypic traits. Stewart (2011) furthers that consistent recording of pedigree information can increase the reliability of evaluations, stating that only limited pedigree information was available for her study.

Stewart (2011) demonstrated interesting trends from Futurity assessed horses. First, mean scores from foals to yearlings decreased, however these peaked again for 3 year old's, suggesting that as the horse matures the assessed traits are easier to evaluate. Second, horses which were entered for showjumping and one other discipline in their Futurity career achieved lower mean scores than those entered for just showjumping, which suggests that discipline specific pre-training can influence horse scores. Lastly, the lower range of the 10 point scoring scale (less than 4) is not used, which may result from two causes; either due to owners pre-selecting horses which match the judging criteria or industry pressures on the judging panel to not excessively penalise a poor performing horse. If judges are effected by industry pressures this can reduce the reliability of the Futurity evaluations. Several undergraduate papers have further identified some interesting Futurity trends. In a Bachelors dissertation by Clausen (2009), veterinary score was demonstrated to have the least influence on overall score; whilst type and temperament had the greatest. In a conference paper by Neyround (2013) it was demonstrated that the highest scoring Futurity horses had significantly lower body condition scores and higher muscular condition scores than those awarded with average body condition and muscular scores. Most recently, in an unpublished Master's thesis by Fisker-Hansen (2015), judge bias was investigated by comparing the Futurity scores of horses with different coat colours. Results demonstrated that lower scores were awarded to horse coat colours of piebald, skewbald and roan, suggesting that some negative judge bias is present in the Futurity scheme. Whilst the discussed papers are not peer reviewed and must therefore be interpreted with caution, it can be argued that the findings demonstrate that further research is needed to better understand the parameters of the Futurity evaluations.

Clearly from all previous research, there is still much to learn regarding the parameters of the Futurity evaluations as horses are scored subjectively on their traits and therefore the resulting premium scores may be influenced by either the personal opinions of judges and by environmental effects. Judge subjectivity has been demonstrated in dressage scoring (Hawson et al. 2010) and has been voiced as a limitation to the Futurity test integrity by British horse breeders (Horse Breeders Magazine 2009). Key aims of the Futurity are to inform better British breeding practices with the use of accurate evaluation results (BEF 2015a). Additionally the Futurity aims to identify potential sport horses for competition performance, either for amateur competition or elite performance, reducing the reliance on internationally importing quality horses, which loses the British Breeding industry money (Kearsley 2008b). If horse premiums are not representative of horse potential the aims of the Futurity will not be met. Furthermore, high horse premium scores have been demonstrated to increase the monetary value of an advertised horse (BEF 2015a), as well as gain them entry to the Futurity Equine Bridge programme (BEF 2014c further details of which can be found in Appendix 1g).

Therefore if horse premiums are an inaccurate representation of horse ability, potential buyers may be misled and entries to the Futurity Bridge will be miss informed. Further research on the parameters of the Futurity can assist in determining if the evaluations are achieving their aims.

Currently the Futurity is capturing a comparable proportion of young horses to international FPT (Stewart 2011). In 2009, the Futurity evaluations attracted 10% (873) of registered British foals (Stewart 2011), and in total from 2007 to 2014, the Futurity database held 5361 records of evaluated horses; however horses can be tested multiple times so the records of individual horses are expected to be less (British Breeding 2015). In comparison, Viklund et al. (2008) used 4,110 horse evaluation records when studying of the parameters of the RHQT. Therefore it is evident that the Futurity database consists of a sufficient number of horse records to be investigated for their reliability. Internationally, FPT reliability is investigated by correlating young horse scores with adult horse competition data (Hellsten et al. 2006). As Stewart's (2011) research suggests the Futurity is predicting positive trait heritabilities from a large dataset, these young horse results can be correlated with competition results. Competition results have demonstrated positive trait heritabilities from British competition databases such as British Eventing (Kearsley et al. 2008a and Stewart et al. 2012), British Showjumping (Stewart 2011) and British Dressage (Stewart et al. 2010). Therefore these competition databases can provide phenotypic information for breeding programme and in addition can be correlated with Futurity scores to assess and further understand the reliability of Futurity evaluations (Kearsley 2008b and Stewart 2011).

The aim of the current study is to investigate the reliability of the horse scores awarded by the BEF Futurity program, as the database is now considered large enough (n=5361 records from 2007 to 2014) to achieve reliable conclusions. To achieve the project aims, retrospective young horse Futurity scores were correlated with adult horse competition results in eventing (BE), showjumping (BS) and dressage (BD) to ascertain if a positive relationship exists between the two measures (Futurity scores and competition score). Horse competition results from BE, BD and BS, were converted into performance measures. Futurity horse scores and component scores were the independent variables. Results from this study highlight the strengths and limitations of the Futurity programme when used to inform the potential of the British sport horses. However due to using retrospective data, it must be noted that the results from this study may not be entirely representative of the current (2015) Futurity system and the wider population of horses.

Chapter 2. Methodology.

Subjects

The present study involved extensive desk-based collection of data from the British Equestrian Federation (BEF) Futurity program and horse competition results from relevant sport governing bodies British Eventing (BE), British Dressage (BD) and British Showjumping (BS). From the BEF Futurity database, horse records from 2007 to 2014 were used of horses born in or before 2010 and evaluated for either eventing, dressage or showjumping. Table 2 summarises the information (data) that each Futurity horse record contained.

Table 2: The information on a Futurity horse record (British Breeding 2015).

The information on a Futurity horse record		
Horse details	**Evaluation details**	**Evaluation score details**
Name	Evaluation date	Average score
Age		Veterinary score
Sex	Venue location	Frame and build score
Coat colour		Walk score
Parentage	Evaluation discipline	Trot score
Registration organisation		Canter score
		Jump score
		Athleticism score

The Futurity program has developed over time and as such, some of the recorded information available is dependent on the year of evaluation. Futurity records from 2007 have no coat colour or dam information. In 2007 and 2008, only average evaluation scores were recorded. Records from 2009 onwards contain all component scores, however the “jump” score is only gained by horses aged over three at the time of evaluation and who were being assessed for the discipline of showjumping or eventing (unless the horse had a foal at foot). In 2007 horse sex was recorded as “Male” or “Female”, this changed to “Gelding”, “Stallion”, or “Mare” from 2008 onwards. For the purpose of this project, horse sex has been simplified to “Male” or “Female” for consistency.

BE, BD and BS all hold comprehensive competition records for their respective disciplines, therefore data from these organisations were used to create performance measures from competition scores. The data used from BE, BD and BS are displayed in Table 3.

Table 3: The data used from BE, BD and BS (British Eventing 2015b, British Dressage 2015 and British Showjumping 2015b).

The data used from BE, BD and BS

British Eventing

Horse details	Competition results
Name	Total points at each grade*
Age	Total foundation points at each grade
Sex	Total placings at each grade
Height	Total penalty points at each grade
Parentage	Total penalty points in each phase**

**Grades of competition; BE80, BE90, BE100, Novice, Intermediate, Advanced.*
***Phases of competition; dressage, cross country, and showjumping.*

British Dressage

Horse details	Competition results
Name	Total placings at each level***
Age	Total points at each level
Sex	
Parentage	Total percentage scored at each level

****Levels of competition; Introductory, prelim, novice, elementary, medium, advanced medium, advanced, FEI Prix St George, FEI Intermediare 1, FEI Intermediare 2, Grand Prix.*

British Showjumping

Horse details	Competition results
Name	Total placings at each class****
Age	Total points at each class
Sex	
Height	Total winnings at each class
Parentage	

*****Classes of competition; 90cm and below, 1m, 1.10m, 1.20m, 1.30m, 1.40m and 1.50m+*

A questionnaire was also distributed to provide feedback from horse owners to establish the competition trends of Futurity evaluated horses. This sought information regarding the following aspects, trends in Futurity attendance and scoring, competing progress of the Futurity horse and public opinions of the usefulness of the evaluations. (See Appendix 2a for the full questionnaire).

Materials

The raw data from the BEF Futurity program was requested through email correspondence with Futurity officials (Dr Jan Rogers and Dr Joanne Dixon). The data received contained all horse evaluations from 2007 to 2014 and was stored on an Excel spreadsheet to undertake this desk based research.

The Microsoft Office programme Excel 2013 was used for collecting and organising all the data for this project. The MiniTab™ v17 statistical package was used to analyse all data.

BE competition data were collected from their public website page; http://www.britisheventing.com/search.asp?section=156&sectionTitle=Search. As BE competition results are publically available, their online database could be manually searched through to obtain data of individual horses competitions records (British Eventing 2015a). Data from BS were collected from their website page; http://www.britishshowjumping.co.uk/ (British Showjumping 2015a). Membership was required for any individual to gain access to BS horse competition results and therefore for the purpose of this study, BS non-jumping yearly membership package was purchased for £54 and with this identification, individual horse competition records were searchable and competition records could be obtained. BD horse competition records were accessed manually on site within their head offices from their internal database of all horses registered and competing with BD. Some unavoidable bias were present in the governing body's records. BE have recorded all results on their system since 1998. BS currently only display competition results of horses which have won prize money or gained a double clear round (no faults gained in their first round and in the jump off). Since 2011 BD record all scores of registered horses, however before this date only scores of 60% and above were recorded.

In additon a questionnaire was designed and published through the website Survey Planet (https://surveyplanet.com/56266b378535760f1470b1d5). The questionnaire was distributed through social media sites, such as Facebook and sent directly to relevant organisations (BEF, BE, BD and BS). The main areas covered in the questionnaire were trends in Futurity attendance and scoring, competing progress of the Futurity horse and public opinions of the usefulness of the evaluations (see Appendix 1a).

BEF Futurity data process

The Futurity master excel document contained 5,359 horse evaluation records. Data of no relevance to the project were discarded. Horses evaluated for Endurance or Sports Pony disciplines were discarded as the data from these disciplines were insufficient for analysis with only n=134 and n=151 records respectively. Futurity records where the name had been recorded as “unknown”, “unnamed” or “foal” were also discarded; this accounted for n=23 of the eventers’ records, n=9 of the dressage records and n=15 of the showjumpers. Any horses born after 2010 were discarded as only records of horses born before 2010 were used to ensure all horses were old enough to compete in either dressage, eventing or showjumping competitions in or before 2014. In BE, BD and BS competitions, horses can compete from the age of four (British Eventing 2015b, British Dressage 2015 and British Showjumping 2015b). The Futurity evaluations commenced in 2007, making the oldest horses in the dataset born in 2003, and the age range of horses from 5 to 11 years old at the time of analysis. Once irrelevant Futurity records were discarded, 3,177 Futurity horse records remained (1,243 evaluated for eventing, 656 for showjumping and 1,276 for dressage). However as horses could be evaluated more than once they could have more than one Futurity record. Actual numbers of horses which attended the Futurity equalled 2,559 (970 for eventing, 540 for Showjumping and 1,049 for Dressage). Therefore it is understood that 20% of horses are evaluated more than once.

Each individual Futurity horse record was match to its corresponding competition record. Competition records were identified by manually searching each evaluated horse name in their respective discipline search engine (BE, BD and BS). The criteria for a match of a Futurity name to a competition record was that the horse name, sex, year of birth and one parent name must match exactly in both records, which gave four measures to confirm the horse identity. In certain cases, Futurity horse names were not spelt exactly the same as the name on the competition record, for example the use of “or” instead of “er” or the use of “ie” instead of “y”. In cases where horse names were very similar to, but not spelt identically on the Futurity record and competition record, stricter measures were applied to confirm the horses’ identity. These stricter measures were that the horse name must be similar and the year of birth, sex, sire and dam must match exactly, giving five measures to confirm identity.

From manual searching of the evaluated horses it was established that n=268 had been registered with BE, n=451 with BD and n=175 with BS, totalling n=894 horses with usable Futurity and competition data for analysis in this project [This data analysis has been included in the method to clearly present what data were available for the main analysing phase]. It must be noted that as a result of some Futurity horse names being spelt differently on

competition records and due to the nature of manually searching the horse names, it was not possible to confirm if all Futurity records were matched with their respective competition records, although every effort was taken to do so.

Competition data process

The competition records of the n=894 Futurity horses were used to create performance measures for analysis. In eventing, dressage and showjumping competitions, the levels of difficulty are progressive, see Table 4 for a scale of competition levels used in this study.

Table 4: The competition levels in BE, BD and BS (British Eventing 2015b, British Dressage 2015 and British Showjumping 2015b).

The competition levels in BE, BD and BS						
British Eventing (Grades)						
BE80(T)	BE90	BE100	Novice	Intermediate	Advanced	
Note: In BE the different phases of competition can be at different grades, but the cross country phase is always kept to the lowest grades and therefore it is this phase that it used to define a competition level.						
British Dressage (Levels)						
Prelim	Novice	Elementary	Medium	Advanced Medium	Advanced	FEI Prix St George	FEI Interdmediare1
Note: BD levels go up to FEI Intermediare 2 and Grand Prix, however no horses in the analysis reach this level and therefore it is not used.						
British Showjumping (Class)						
90cm and below	1m	1.10m	1.20m	1,30m	1.40m	1.50m+
Note: BS class heights go up in increments of 5cm, however classes were group to increments of 10cm to prevent low numbers of horses in the subgroups.						

In some cases in BE and BS competition results, competition levels were recorded under Fédération Equestre Internationale (FEI) starred terms from levels 1 to 4. Where FEI starred terms were used in eventing, the following rule applied; 1star was equivalent to Novice, 2star was equivalent to Intermediate, 3star was equivalent to Intermediate and 4star was equivalent to Advanced (British Eventing 2015b). In showjumping, FEI starred terms can only be approximately converted into class heights. For example, a 1star class is defined as any course less than 1.40m (Fédération Equestre Internationale 2014). Therefore to increase accuracy, where FEI starred terms were used in records, the specific competition name, venue and competition date were searched online to assign the competition to the correct height group. In addition, BS assigns class names to the heights of competition, however for the

purpose and simplicity of this research investigation all class names were converted to their respective heights (British Showjumping 2015b). If any competition result was incomplete on a horse's record, i.e. missing the level or placing information, or in the case of eventing not completing all three phases of competition, then it was excluded from the analysis. The measures of performance used in this study are summarised in Table 3 and discussed in more detail below.

Performance measure 1 – Horses' number of competition levels competed in

BE, BD and BS all record the competition levels which a horse has competed in. The highest level each individual horse competed in was converted into a numerical value and used in this study. For example, a horse competing at only BE80(T), BD Preliminary or BS 90cm were scored as a 1, whereas a horse competing at BE Advanced BD Advanced or BS 1.40m were scored as a 6.

Performance measure 2 - Horse placings

BE, BD and BS all record horse competition placings. All individual horse placings were totalled and averaged for analysis (removing the variable of the number of competitions a horse competed in). Horse placings were used in this study as all horses could be evaluated by this measure and have been used in previous studies to evaluate horse performance (Braam et al. 2011, Viklund et al. 2010a and Olsson et al. 2008). However horse placings are influenced by external factors such as the number of competitors and their ability, which then influences the measured horses score, reducing integrity.

Performance measure 3 – Horse points

BE, BD, and BS all record horse competition points; these were totalled and averaged for analysis. In BE points are awarded at Novice level and above. Foundation points are awarded at B90 and BE100 levels and were only introduced in 2010, making this measure less comprehensive than points. The allocation of BE points depends on the horse's placing, grade of competition and the number of starters in the dressage phase (British Eventing 2015b). In BD points are awarded at Novice level and above. Prelim points are awarded in the restricted sections, but not the open sections. The restricted section is open to those who have less than 14 points at the next level up from which they are competing (British Dressage 2015). In BS points are awarded based on placing and the class of competition. From 14.04.14, BS replaced their previous version of points (notional winnings) with the points system used today (British Showjumping 2014a). Previous notional winnings could be converted into points from the conversion of £1 = 1point, therefore points and notional winnings have been used as the same measure in this study.

In BE and BS, points are influenced by the ability and number of competitors; not all horses who compete will earn points, therefore not all horses can be analysed by this method. However points have been used in this study as once a horse reaches the point's threshold of a level of competition in BE, BD and BS, they must progress to the next to level, therefore the number of points a horse has represents their ability. Researchers Olsson et al. (2008), Hellsten et al. (2006) and Wallin et al. (2003) all support the use of points to measure horse competition ability.

Performance measure 4 – Horse monetary winnings

In BS competition records, horse actual monetary winnings are recorded and were totalled and averaged for analysis in this study. Horse actual winnings is a record of the full monetary value which a horse has won. The amount of winnings earnt at a competition depends on the level, the number of competitors, the cost of the entry fee, and the total prize fund available, therefore actual winnings are not a direct measure of an individual horse's performance. In addition, not all horses which compete earn winnings. Nevertheless more money is awarded to the highest placing horses and therefore this method has been used as a measure of performance in this study. Furthermore, researchers Langlois and Blouin (2004) and Ricard and Chanu (2001) support the use of using horse earnings to evaluate performance.

Performance measure 5 – Horse percentage score

In BD competition records, horse competition scores are recorded by percentages; the higher the percentage, the better the performance. All horse percentages were totalled and averaged for analysis. As all horses which compete earn a percentage score, all horses in this study can be analysed. The use of BD percentage scores to evaluate dressage horse performance is supported in research by Stewart et al. (2010).

Performance measure 6 – Horse penalty points

In BE competition records, horse penalty points are recorded for all competing horses and were totalled and averaged in this study. In BE competitions the horse competes through three phases (dressage, showjumping and cross country) and penalty points are recorded for each phase. By combining phase of competition with penalty points, the eventing horses' performance could be analysed at each individual phase. Penalty points are not influenced by the ability of competitors, therefore only the performance of the individual is measured. The use of BE penalty points as a method of analysis is supported by Kearsley et al. (2008b) and Stewart et al. (2012) who used penalty points to evaluate eventing horse performance.

Further information of how points, winnings, percentages and penalty points are allocated can be found in the governing bodies sport handbook (British Eventing 2015b, British Dressage 2015 and British Showjumping 2015b).

Each performance measure was calculated at each level of competition, resulting in 34 traits for analysis for eventing horses, 32 traits for dressage horses and 28 traits for showjumping horses (see Table 5).

Table 5: The traits available for analysis from BE, BD and BS competition records (British Eventing 2015b, British Dressage 2015 and British Showjumping 2015b).

The traits available for analysis from BE, BD and BS competition records						
British Eventing	Placing	Points	Penalty points			
			Dressage	Cross Country	Showjumping	Overall
BE80T						
BE90						
Novice						
Intermediate						
Advanced						

29 traits for analysis

British Dressage	Placing	Points	Percent
Prelim			
Novice			
Elementary			
Medium			
Advanced Medium			
Advanced			
FEI Prix St George			
FEI Intermediare 1			

24 traits for analysis

British Showjumping	Placing	Points	Actual winnings
90cm and below			
1m			
1.10m			
1.20m			
1.30m			
1.40m			
1.50m+			

21 traits for analysis

Dates of data collection

As horse competition results are continuously updating, dates of data collection were put in place to keep to time constraints of the project. The specific dates of data collection are presented in Table 6. Any horses registered and any competition results accrued after the dates of data collection were not used in this study. Data from the BEF Futurity questionnaire were collected from 20.10.15 to 20.1.16, giving a total of 91 days of data collection and 59 respondents.

Table 6: Dates of data collection.

Dates of data collection		
Dates of data collection	**Registered horses collected before**	**Competition results collected before**
British Eventing	07.06.15	25.08.15
British Dressage	19.11.15	20.11.15
British Showjumping	23.07.15	10.08.15

Analysis

Horse Futurity records were paired with their corresponding competition results and split into three discipline specific spreadsheets for analysis (eventing horses', dressage horses' and showjumping horses'). Futurity evaluation scores (the overall score and component scores) were summed together and divided by the number of numbers used to create an average score (for horses which had more than one Futurity score to their name). For horses which had only attended one Futurity evaluation, average scores was not required as the single overall and component scores were used. Horse competition performance scores followed a similar process. All relevant performance scores were summed together and divided by the number of numbers used to create an average score. For example, for a horse which gained 3 placings results at BE100, these placings were summed together and divided by 3 to create an average score. Averaging horse scores made them more comparable to each other as the variable of the number of evaluations/competitions a horse attended was removed. Data distributions/trends of the Futurity data and questionnaire results were established by calculating means, ranges and percentages of each evaluation group (eventing, dressage and showjumping). General regression analysis was used to establish if the dependant variable of Futurity score were influenced by the independent variables of horse age, sex, coat colour and evaluation location. Finally regression analysis was used to investigate any relationship between averaged Futurity and competition scores in each discipline specific group (eventing, dressage and showjumping). All data were transferred into Minitab v17 statistical software for analysis.

Chapter 3. Results.

Distribution statistics

Futurity horses with competition results

Of all horses evaluated by the Futurity, n=2,559 were born between 2003 and 2010, n=894 (34.9%) of these had been registered with an equestrian governing body and n=566 (22.1%) had achieved competition results. It is these n=566 horse with competition results which are subject to analysis in this study and they will be referred to as the Futurity group throughout.

The full breakdown of Futurity horses with competition results is displayed in Table 7 and the percentage distributions are in Table 8**.**

Table 7. The distribution of Futurity horses used in this study.

The number of Futurity horses used in this study			
Futurity data distribution	**Number of horses born between 2003 to 2010**	**Number of horses registered with a governing body**	**Number of horses with competition results**
Total	2,559	894	566
Eventing	970	268	182
Showjumping	540	175	127
Dressage	1,049	451	257

Table 8. The data distribution of Futurity horses available for this study (percentages).

The distribution of Futurity horses used in this study (percentages)				
Futurity data distribution	**Horses born between 2003 to 2010**	**Horses registered with a governing body**	**Horses with competition results**	**Horses with competition results**
Total	2,559	34% of total	22% of total horses	63% of registered horses
Eventing	37% of total	27% of eventing total	10% of total horses	67% of registered horses
Showjumping	21% of total	32% of showjumping total	23% of total horses	72% of registered horses
Dressage	40% of total	42% of dressage total	24% of total horses	56% of registered horses

Distributions of age of the Futurity group

In the dressage and eventing groups, horse year of births ranged from 2003 to 2010. In the showjumping group, the range was 2004 to 2010. Horse competition data were collected in 2015, therefore horse age in competition ranged from 5 to 12 years old. Mean year of birth of the dressage group equalled 2008, the eventing group 2007, and the showjumping group 2009. The horse age at evaluation can vary from foal to three year old. In the dressage (n=257) and showjumping (n=127) groups, the majority of horses were assessed as foals (30.3% and 31.4% respectively), in the eventing (n=182) group most were assessed as three year olds (30.2%). Full horse age at evaluation are displayed in Figure 1 and the numerical distributions in Table 9**.**

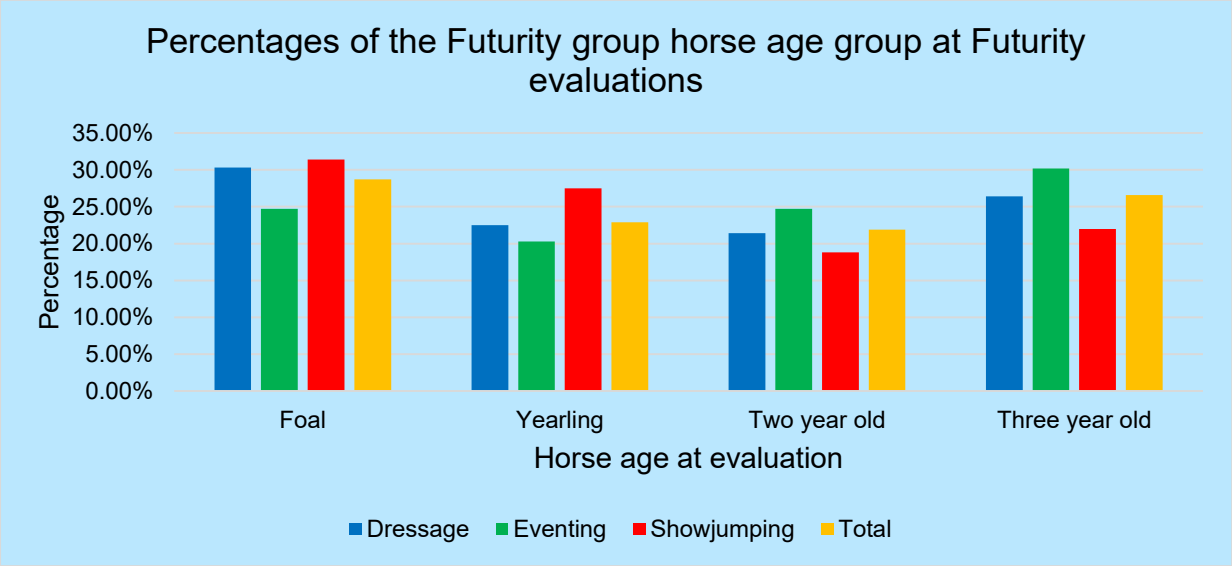


Figure1. Percentages of the Futurity group horse age group at Futurity evaluations.

Table 9. Number of Futurity group horse age groups at Futurity evaluations.

Number of Futurity group horse age groups at Futurity evaluations				
	Foal	**Yearling**	**Two year old**	**Three year old**
Dressage	78	58	55	68
Eventing	45	37	45	55
Showjumping	40	35	24	28
Total	163	130	124	151

Distributions of sex of the Futurity group

Horses were classed as male or female. In the dressage group, 51.3% (132/257) of horses were male, in eventing 60.9% (111/182) were male, and 61.4% (78/127) were male in showjumping. Full distribution of horse sexes is summarised in Table 10**.**

Table 10. Futurity group horse sexes at Futurity evaluations.

Futurity group horse sexes at Futurity evaluations		
	Male	**Female**
Dressage	132	125
Eventing	111	71
Showjumping	78	49
Total	321	245

Distributions of coat colour of the Futurity group

Coat colour was recorded by the Futurity from 2008 onwards. Table 11 displays the distribution of the recorded horse coat colour at Futurity evaluations.

Table 11. Futurity group horse coat colours recorded at Futurity evaluations.

Futurity group horse coat colours recorded at Futurity evaluations											
	Bay	**Light bay**	**Dark bay**	**Black**	**Chestnut**	**Dark chestnut**	**Grey**	**Piebald**	**Skewbald**	**Roan**	**Palomino**
Dressage	92	4	39	40	40	5	3	1	6	1	4
Eventing	66	2	33	7	27	1	12	0	9	1	0
Showjumping	56	6	6	6	23	0	21	2	3	2	2
Total	**214**	**12**	**78**	**53**	**90**	**6**	**36**	**3**	**18**	**4**	**6**

Distributions of horse premiums of the Futurity group

Horse scores at the Futurity are awarded on a 10 point scale. In the dressage group, scores awarded ranged from 6.9 to 9.9, scores in the eventing group ranged from 6.6 to 9.7, and in the showjumping group 5.6 to 9.8. The mean score awarded in the dressage and showjumping groups were both 8.3, whilst in eventing this was 8.2. In the dressage group, first and higher first premiums were most commonly awarded and equal 34.7% (89/257) each. In the eventing and showjumping groups first premiums were awarded to 40.8% (74/182) and 46% (59/127) of horses respectively. Full percentage distributions are displayed in Figure 2 and full details are in Appendix 3a.

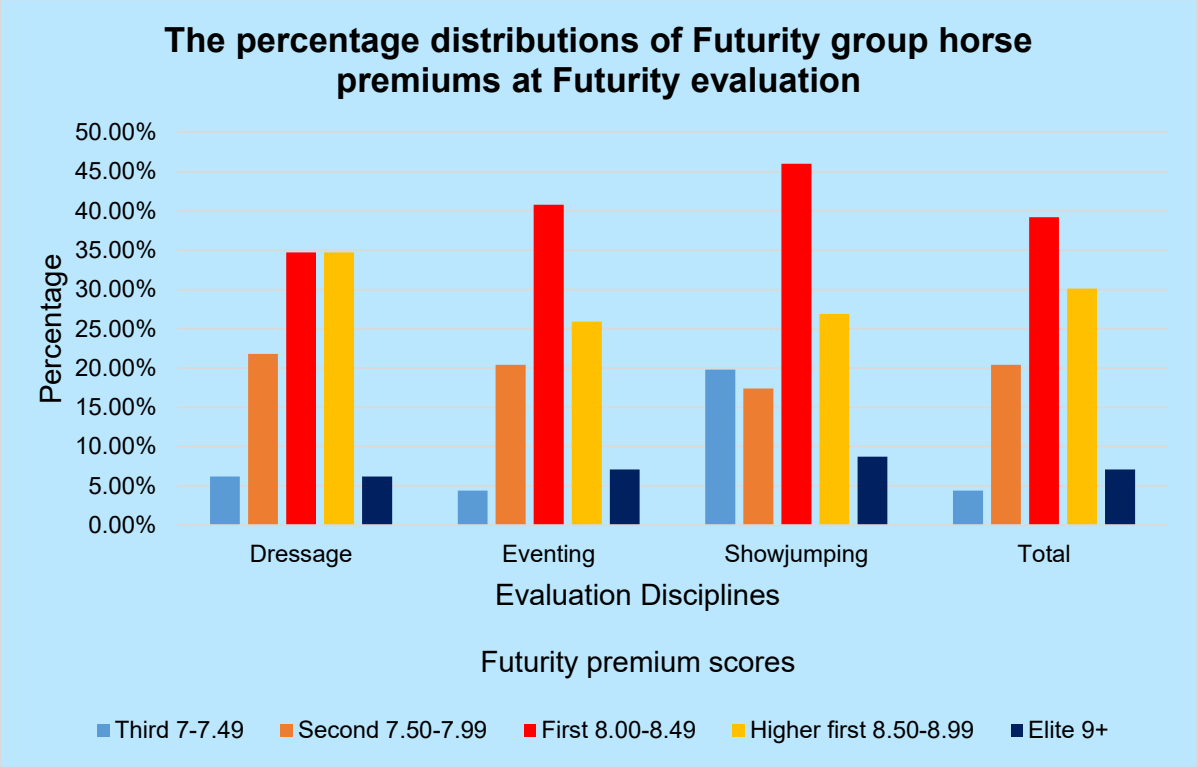


Figure 2. The percentage distributions of Futurity group horse premiums at Futurity evaluation

Distributions of evaluation location of the Futurity group

Futurity evaluations take place across England. Table 12 displays the frequency use of evaluation locations of the horses used in this study.

Table 12. Futurity group horse evaluation locations.

Futurity group horse evaluation locations				
Locations	**Dressage**	**Eventing**	**Showjumping**	**Total**
Alsager Equestrian Centre Stoke-on-Trent	1	1	0	2
Arena UK Grantham	24	1	11	36
Catherston Stud Hants	38	20	3	61
College EC Bedfordshire	15	12	9	36
Crofton Manor Hampshire	3	1	0	4
Fountain Equestrian Centre Aberdeen	11	9	2	22
Hartpury College, Gloucestershire	3	0	1	4
Heart of England Stone Staffordshire	11	10	6	27
Myerscough College Preston Lancashire	22	7	8	37
Osbaldeston	4	6	1	11
Plumpton College Lewes East Sussex	21	12	13	46
Richmond Equestrian Centre North Yorkshire	12	10	10	32
Scottish National Equestrian Centre, Edinburgh	1	0	1	2
Solihull RC West Midlands	30	27	31	88
Southview EC, Cheshire	4	2	5	11
Sunnybank EC Rudry Caerphilly	14	4	3	21
Tall Trees Cornwall	7	5	6	18
The Grange Okehampton Devon	14	22	16	52
Writtle College Essex	14	2	1	17

Distributions of component scores of the Futurity group

Futurity horses evaluated from 2008 onwards had component scores recorded. Table 13 summarises the number of Futurity group horses with component scores available for analysis.

Table 13.The number of Futurity group horses with component scores

The number of Futurity group horses with component scores							
	Veterinary mark	**Frame and build**	**Walk**	**Trot**	**Canter**	**Jump**	**Athleticism**
Dressage	180	180	180	180	37	0	180
Eventing	249	249	249	249	81	42	249
Showjumping	135	135	135	135	26	20	135

Distributions of Futurity group horse's competition data

All Futurity group horses have to meet the criterion of competed in one class of competition in this study. Figure 3 displays the competing trends of Futurity group horses. Horse birth years have been grouped in Figure 4 to display the number of competition results available in each birth group. Appendix 3b displays the number of competition results available in each birth grouped and for each discipline. Further details of the trends present in each discipline are displayed in Appendix 3c.

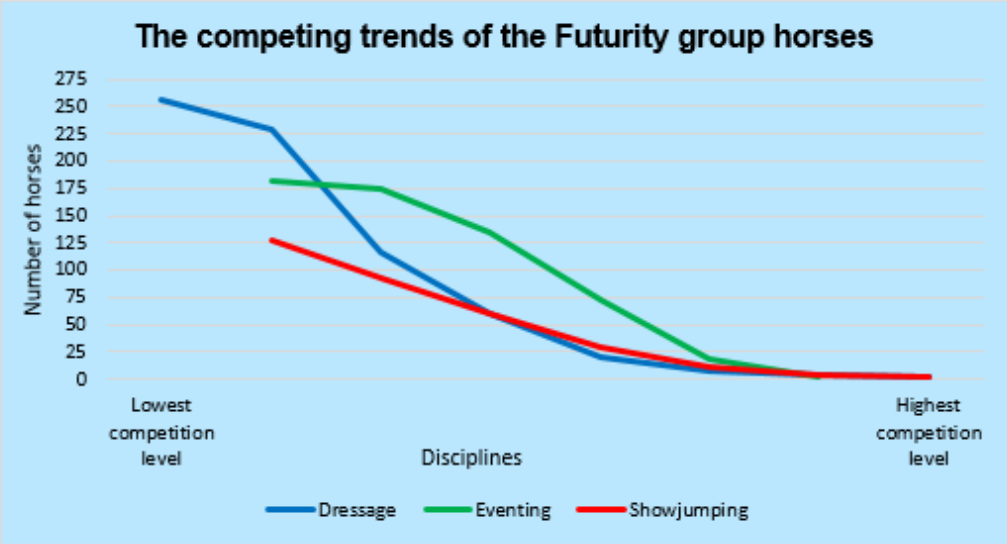


Figure 3. The competing trends of the Futurity group horses.

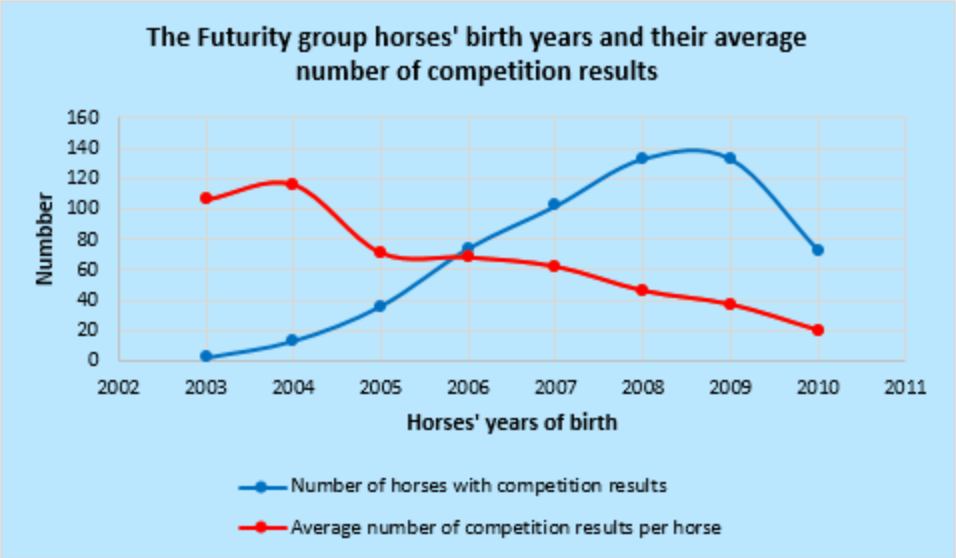


Figure 4. The Futurity group horses' birth years and their average number of competition results.

Regression analysis of the Futurity group

Futurity variables of horse sex, age at evaluation, coat colour and evaluation location were used as the predictor variables to the response of Futurity evaluation score of the Futurity group by regression testing. A number of regression analyses were conducted.

Regression analysis: Average Futurity score versus sex

Horse sex did not statistically predict average Futurity scores in any discipline groups (all regressions non-significant P>0.05).

Regression analysis: Average Futurity score versus age at evaluation

Horse age at evaluation significantly predicted average Futurity scores in the showjumping group ($F_{1,125}$ = 5.87, P<0.05). Horse age at evaluation accounted for 4.5% (R^2) of the explained variability in average Futurity score (R^2adj=3.7%). The regression equation is Fscore=8.22+0.109(Year at evaluation). A statistically significant relationship between horse age at evaluation and average Futurity score was not found in either the dressage or eventing groups (P>0.05).

Regression analysis: Average Futurity score versus coat colour

Horse coat colour significantly predicted average Futurity scores in the eventing group ($F_{1,161}$=4.97, P<0.05). Horse coat colour accounted for 3% (R^2) of the explained variability in average Futurity score (R^2adj=2.4%). The regression equation is Fscore=8.43-0.0340(Colour). A statistically significant relationship between horse coat colour and average Futurity score was not found in either the dressage and showjumping groups (P>0.05).

Regression analysis: Average Futurity score versus location

Evaluation location not statistically predict average Futurity scores in any discipline groups (all regressions non-significant P>0.05).

Regression analysis of Futurity groups evaluation scores versus competition data, grouped by year of birth

Futurity group horse's average evaluation score were used as the predictor variable to the response variables of horse averaged performance measures (see Table 5 for performance measures). General regression analysis was carried out on each competition discipline with each competition level. Horse year of birth was used as the category predictor to group the horses by birth year. Regression probability value was used to establish significance ($P<0.05$). Only the statistically significant relationships between Futurity scores and the performance measures are illustrated in Table 14. Futurity scores which did not statistically significantly predict any performance scores are not included.

Table 14. Futurity group regression analysis: Performance measure (averaged) versus Futurity score (averaged), Year of birth.

Regression analysis of the Futurity group

Test: Averaged* performance measure (dependent variable) versus averaged* Futurity score (independent variable), Year of birth (categorical predictor)

Performance measures	Regression coefficient	Standard error	P-value (associated with the regression coefficient	Cases used	DF** for error
Percentage					
BD medium	1.9911	0.97882	0.047	61	53
Penalty points (Dressage phase)					
BE90	-2.0742	0.74642	0.006	157	148
BE100	1.4740	0.59128	0.014	131	122
BE Novice	-2.7098	1.2371	0.032	67	59
Penalty points (Showjumping phase)					
BE90	1.54058	0.59542	0.011	157	148

*Averaged scores were calculated by summing the relevant values then dividing by the number of numbers used.
**Degrees of freedom.

Table 14 interpretation

When horse Futurity scores increase by 1point, BD percentage scores at medium level increased by 1.9911 (see figure 5). Therefore an increase in horse performance scoring at Futurity increases the percentage scores earnt in BD medium classes. The higher the percentage earnt in BD, the better the horse performed, making this an expected result. Average horse Futurity scores accounted for 11.12% (R^2) of the explained variability in percentage scores earnt in BD medium classes (R^2adj=-0.62%).

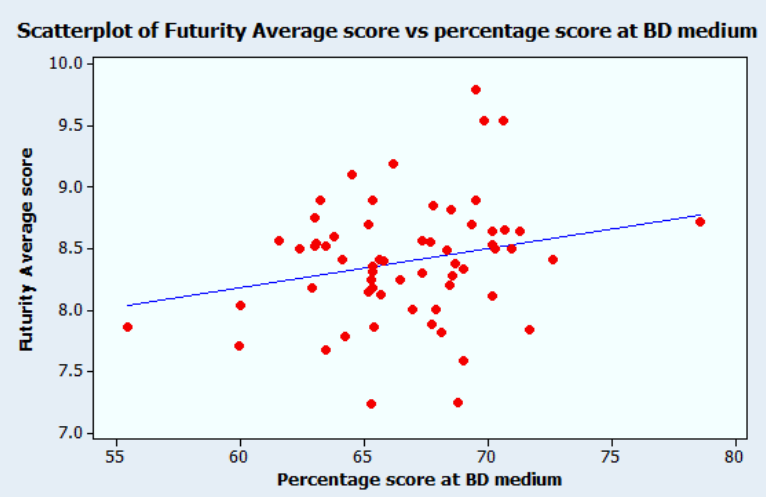


Figure 5. Scatterplot of Futurity average score vs percentage score at BD medium.

When horse Futurity scores increase by 1point, BE90 dressage penalty points scores decrease by 2.074 (see figure 6). Therefore an increase in horse performance scoring at Futurity decreases the penalty point's earnt in the BE90 dressage phase. The lower the penalty points earnt in BE, the better the horse performed, making this an expected result. Average horse Futurity scores accounted for 8.55% (R^2) of the explained variability in percentage scores earnt in BD medium classes (R^2adj=3.61%).

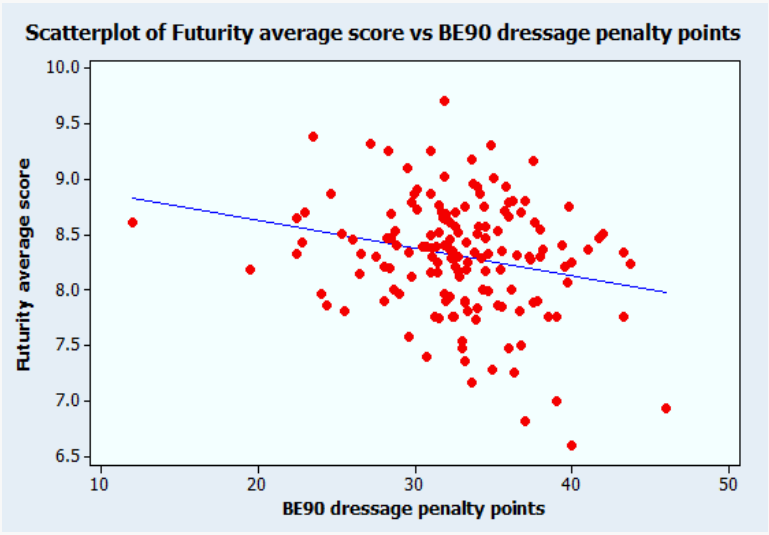


Figure 6. Scatterplot of Futurity average score vs BE90 dressage penalty points.

When horse Futurity scores increase by 1point, BE100 dressage penalty points scores increase by 1.4740 (see figure 7). Therefore an increase in horse performance scoring at Futurity increases the penalty point's earnt in the BE100 dressage phase. The higher the penalty points earnt in BE, the worse the horse performed, making this an unexpected result. Average horse Futurity scores accounted for 16.63% (R^2) of the explained variability in percentage scores earnt in BD medium classes (R^2adj=11.16%).

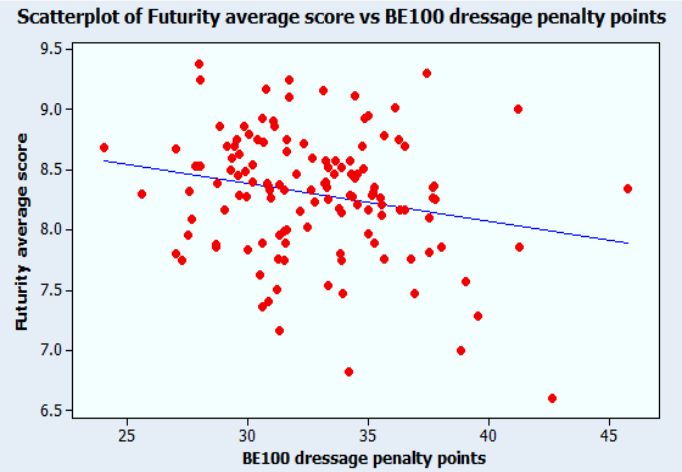


Figure 7. Scatterplot of Futurity average score vs BE100 dressage penalty points.

When horse Futurity scores increase by 1point, BE novice dressage penalty point scores decrease by 2.7098 (see figure 8). Therefore an increase in horse performance scoring at Futurity decreases the penalty point's earnt in the BE novice dressage phase. The lower the penalty points earnt in BE, the better the horse performed, making this an expected result. Average horse Futurity scores accounted for 18.41% (R^2) of the explained variability in percentage scores earnt in BD medium classes (R^2adj=8.73%).

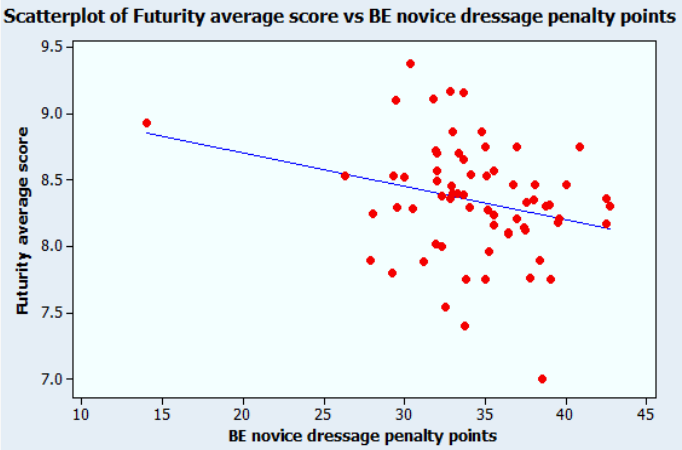


Figure 8. Scatterplot of Futurity average score vs BE novice dressage penalty points.

When horse Futurity scores increase by 1point, BE90 showjumping penalty point scores increase by 1.54058 (see figure 9). Therefore an increase in horse performance scoring at Futurity increases the penalty point's earnt in the BE90 showjumping phase. The higher the penalty points earnt in BE, the worse the horse performed, making this an unexpected result. Average horse Futurity scores accounted for 9.90% (R^2) of the explained variability in percentage scores earnt in BD medium classes (R^2adj=5.03%).

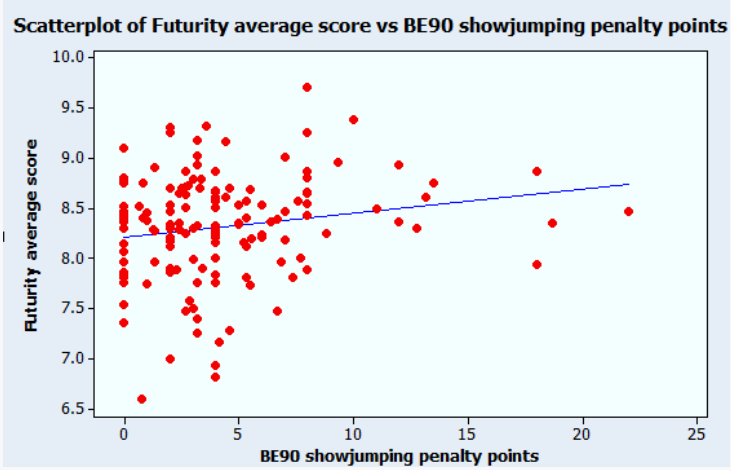


Figure 9. Scatterplot of Futurity average score vs BE90 showjumping penalty points.

Table 14 displays positive correlations between average Futurity scores and horse performance measures. Futurity scores correlated with 1 BD performance measure; percentage scores. An increase in Futurity score increased BD percentage scores, which is an expected result as a higher Futurity score correlated with better horse performance at competition. However only 61 cases were used. Futurity scores correlated with 4 BE performance measures. 2 of these were expected results as a higher Futurity score correlated with better horse performance at competition. However 2 were unexpected as a higher Futurity score correlated with worsened competition performance. BE penalty points in dressage were the measure most correlated with Futurity scores. Futurity scores correlated with none BS performance measures. For all positive Futurity and performance correlations, R^2adj values ranged from 3.61% to 11.16% suggesting that the model only explains a small proportion of the variability of the response data, other factors can be responsible.

Regression analysis of Futurity groups evaluation component scores versus competition data, grouped by year of birth

Futurity group horse's average evaluation component scores were used as the predictor variable to the response variables of horse averaged performance measures (see Table 5 for performance measures). Regression analysis was carried out on each competition discipline and level, horse year of birth was used as the category predictor to group the horses by birth year and the regression tests probability value was used to establish significance ($P<0.05$). Only the statistically significant relationships between Futurity component scores and the performance measures are illustrated in Table 15. Futurity component scores which did not statistically significantly predict any performance scores are not included.

Table 15. Futurity group regression analysis: Performance measure (averaged) versus Futurity component scores (averaged), year of birth.

Regression analysis of the Futurity group

Test: Averaged* performance measure (dependant variable) versus averaged* Futurity component scores (independent variable), Year of birth (categorical predictor)

Performance measures	Regression coefficient	Standard error	P-value (associated with the regression coefficient)	Cases used	DF for error
Component score: Veterinary mark					
Points					
BD medium	1.44036	0.65275	0.036	32	27
BS 1.20m	5.2548	2.1895	0.030	21	15
BS 1.30m	14.089	3.8647	0.036	8	3
Winnings					
BS 90cm	9.6416	4.1639	0.023	95	88
Penalty points (overall)					
BE novice	10.9847	4.3754	0.016	43	38
Penalty points (showjumping)					
BE100	-2.2347	0.83678	0.009	90	84
Penalty points (cross country)					
BE novice	8.9717	3.6158	0.018	43	38
Component score: Frame and build					
Percentage					
BD advanced	-12.771	0.31342	0.016	4	1
Points					
BD novice	0.386853	0.14761	0.010	151	145
BD advanced	-2.1276	0.80079	0.045	10	5
Winnings					
BS 90cm	10.0877	4.6973	0.034	95	88
Penalty points (dressage)					
BE90	-1.5200	0.73579	0.041	116	110
Component score: Walk					
Percentage					
BD medium	2.8477	1.2044	0.026	32	27
BD advanced	-4.677	0.31088	0.042	4	1
Points					
BD novice	0.68866	0.20898	0.001	151	145
BD medium	1.6105	0.56186	0.008	32	27
Winnings					
BS 1.30m	30.394	6.6324	0.020	8	3
Penalty points (dressage)					

BE90	-2.0342	0.84279	0.017	116	110
Penalty points (showjumping)					
BE90	1.60111	0.68875	0.022	116	110
Component score: Trot					
Percentage					
BD advanced	-4.257	0.104473	0.016	4	1
Points					
BD novice	0.60760	0.17652	0.001	151	145
Placings					
BS 1.20m	-2.2529	1.03254	0.045	21	15
Winnings					
BS 1.30m	24.643	5.0444	0.016	8	3
Penalty points (dressage)					
BE novice	-3.5017	1.4478	0.020	43	38
Component score: Canter					
Percentage					
BD novice	2.9634	1.19749	0.019	32	29
Points					
BD novice	1.29463	0.46792	0.010	31	28
Penalty points (cross country)					
BE novice	-2.8888	0.14133	0.031	3	1
Component score: Jump					
Penalty points (showjumping)					
BE100	5.1757	1.4420	0.006	12	9
Component score: Athleticism					
Percentage					
BD elementary	0.7976	0.34634	0.024	72	66
BD advanced	-2.338	0.15544	0.015	4	1
Points					
BD novice	0.34508	0.09456	<0.0001	151	145
BD elementary	0.35600	0.15732	0.027	73	67
BD medium	0.63052	0.26643	0.025	32	27
Penalty points (dressage)					
BE novice	-1.3501	0.57133	0.023	43	38
Penalty points (showjumping)					
BE100	-1.0802	0.29289	<0.0001	90	84

*Averaged scores were calculated by summing the relevant values then dividing by the number of numbers used
**Degrees of freedom

Table 15 interpretation

Component scores vs British Dressage performance measures

Veterinary component score:

When horse Futurity veterinary scores increase by 1point, BD point scores at medium level increase by 1.44036. Therefore an increase in veterinary scoring at the Futurity increases the point's earnt in BD medium classes. The higher the point's earnt in BD, the better the horse performed, making this an expected result.

Frame and build component score:

When horse Futurity frame and build scores increase by 1point, BD percentage scores at advanced level decrease by 12.771 and BD points scores at advanced level decrease by -2.1276. Therefore an increase in frame and build scoring at the Futurity decreases the

percentage and point scores earnt in BD advanced classes. The lower the percentage and point scores earnt in BD, the worse the horse performed, making these unexpected results. However, when horse Futurity frame and build scores increase by 1point, BD point scores at novice level increase by 0.386853. The higher the point's earnt in BD, the better the horse performed, making this an expected result.

Walk component score:

When horse Futurity walk scores increase by 1point, BD percentage scores at medium level increase by 2.8477 and BD point scores at novice and medium level increase by 0.386853 and 1.6105 respectively. Therefore an increase in walk scoring at the Futurity increases the point's earnt in BD novice and medium classes and percentage scores at medium level. The higher the percent and point scores earnt in BD, the better the horse performed, making these expected results. However, when horse Futurity walk scores increase by 1point, BD percentage scores at advanced level decrease by -4.677. The lower the percentage scores earnt in BD, the worse the horse performed, making this an unexpected result.

Trot component score:

When horse Futurity trot scores increase by 1point, BD point scores earnt at novice level increase by 0.60760. Therefore an increase in trot scoring at the Futurity increases the point scores earnt in BD novice level. The higher the point's earnt in BD, the better the horse performed, making this an expected result. However, when horse Futurity trot scores increase by 1point, BD percentage scores earnt at advanced level decrease by -4.257. The lower the percentage earnt in BD, the worse the horse performed, making this an unexpected result.

Canter component score:

When horse Futurity canter scores increase by 1point, BD percentage scores earnt at novice level increase by 2.9634 and BD point scores earnt at novice level increase by 1.29463. Therefore an increase in canter scoring at the Futurity increases the point and percent scores earnt in BD novice level. The higher the percent and point scores earnt in BD, the better the horse performed, making these expected results.

Athleticism component score:

When horse Futurity athleticism scores increase by 1point, BD percentage scores earnt at elementary level increase by 0.7976, BD points scores earnt at novice, elementary and medium level increase by 0.34508, 0.35600 and 0.63052 respectively. Therefore an increase in athleticism scoring at the Futurity increases the percent scores earnt at elementary level and the point scores earnt at novice, elementary and medium level. The higher the percent

and point scores earnt in BD, the better the horse performed, making these expected results. However, when horse Futurity athleticism scores increase by 1point, BD percentage scores earnt at advanced level decrease by 2.338. The lower the percentage earnt in BD, the worse the horse performed, making this an unexpected result.

Component scores vs British Eventing performance measures

Veterinary component score:

When horse Futurity veterinary scores increase by 1point, BE penalty point scores at novice level overall increase by 10.9847 and BE penalty point scores at novice level in the cross country phase increase by 8.9717. Therefore an increase in veterinary scoring at the Futurity increases the penalty points earnt in the BE novice overall competition and in the cross country phase individually. The higher the penalty points earnt in BE, the worse the horse performed, making these unexpected results. However, when horse Futurity veterinary scores increase by 1point, BE100 penalty point scores in the showjumping phase decrease by 2.2347. The lower the penalty points earnt in BE, the better the horse performed, making this an expected result

Frame and build component score:

When horse Futurity frame and build scores increase by 1point, BE90 penalty point scores earnt in the dressage phase decrease by 1.5200. Therefore an increase in frame and build scoring at the Futurity decreases the penalty point's earnt in the BE90 dressage phase. The lower the penalty points earnt in BE, the better the horse performed, making this an expected result.

Walk component score:

When horse Futurity walk scores increase by 1point, BE90 penalty points earnt in the dressage phase decrease by 2.0342. Therefore an increase in walk scoring at the Futurity decreases the penalty point's earnt in the BE90 dressage phase. The lower the penalty points earnt in BE, the better the horse performed, making this an expected result. However, when horse Futurity walk scores increase by 1point, BE90 penalty points earnt in the showjumping phase increase by 1.60111. The higher the penalty points earnt in BE, the worse the horse performed, making this an unexpected result.

Trot component score:

When horse Futurity tort scores increase by 1point, BE novice penalty point's earnt in the dressage phase decrease by 3.5017. Therefore an increase in trot scoring at the Futurity

decreases the penalty point's earnt in the BE novice dressage phase. The lower the penalty points earnt in BE, the better the horse performed, making this an expected result.

Canter component score:

When horse Futurity canter scores increase by 1point, BE novice penalty point's earnt in the cross country phase decrease by 2.8888. Therefore an increase in canter scoring at the Futurity decreases the penalty point's earnt in the BE novice cross country phase. The lower the penalty points earnt in BE, the better the horse performed, making this an expected result.

Jump component score:

When horse Futurity jump scores increase by 1point, BE100 penalty point's earnt in the showjumping phase increase by 5.1757. Therefore an increase in jump scoring at the Futurity increases the penalty point's earnt in the BE100 showjumping phase. The higher the penalty points earnt in BE, the worse the horse performed, making this an unexpected result.

Athleticism component score:

When horse Futurity athleticism scores increase by 1point, BE novice penalty point's earnt in the dressage phase decrease by 1.3501 and BE100 penalty point's earnt in the showjumping phase decrease by 1.0802. Therefore an increase in athleticism scoring at the Futurity decreases the penalty point's earnt in the BE novice dressage phase and BE100 showjumping phase. The lower the penalty points earnt in BE, the better the horse performed, making these expected results.

Component scores vs British Showjumping performance measures

Veterinary component score:

When horse Futurity veterinary scores increase by 1point, BS winning scores earnt in 90cm classes increase by 9.6416 and BS point scores in 1.20m and 1.30m increase by 5.2548 and 14.089 respectively. Therefore an increase in veterinary scoring at the Futurity increases the winnings earnt in BS 90cm classes and the point's earnt in 1.20m and 1.30m classes. The higher the winnings and points earnt in BS, the better the horse performed, making this an expected result.

Frame and build component score:

When horse Futurity frame and build scores increase by 1point, BS winnings earnt in 90cm classes increase by 10.0877. Therefore an increase in frame and build scoring at the Futurity increases the winnings earnt in BS 90cm classes. The higher the winnings earnt in BS, the better the horse performed, making this an expected result.

Walk component score:

When horse Futurity walk scores increase by 1point, BS winnings earnt in 1.30m classes increase by 30.394. Therefore an increase in walk scoring at the Futurity increases the winnings earnt in BS 1.30m classes. The higher the winnings earnt in BS, the better the horse performed, making this an expected result.

Trot component score:

When horse Futurity trot scores increase by 1point, BS winnings earnt in 1.30m class's increase by 24.643 and BS placings reached in 1.20m class's decrease by 2.2529. Therefore an increase in trot scoring at the Futurity increases the winnings earnt in BS 1.30m classes and decreases the placing reached in BS 1.20m classes. The higher the winnings earnt and the lower the placing in BS, the better the horse performed, making these expected results.

Table 15 summary

Table 15 illustrates positive correlations between Futurity component scores and horse performance measures. Futurity components scores correlated with 17 BD performance measures. 12 of these were expected results as a higher Futurity score correlated with better horse performance at competition. However 5 were unexpected as a higher Futurity score correlated with worsened competition performance. BD points was the measure most correlated with Futurity scores and the score of Athleticism correlated the most with BD performance measures.

Futurity components scores correlated with 11 BE performance measures. 7 of these were expected results as a higher Futurity score correlated with better horse performance at competition. However 4 were unexpected as a higher Futurity score correlated with worsened competition performance. BE penalty points in dressage and showjumping were the measure most correlated with Futurity scores and the score of veterinary correlated the most with BE performance measures.

Futurity components scores correlated with 7 BS performance measures. 7 of these were expected results as a higher Futurity score correlated with better horse performance at competition. BS points were the measure most correlated with Futurity scores and the score of veterinary correlated the most with BS performance measures.

Scatterplots and the coefficient of determination for Table 15 can be found in Appendix 3c and 3d respectively.

Questionnaire results

Data from the BEF Futurity questionnaire were collected from 20.10.15 to 20.1.16, giving a total of 91 days for data collection and 59/72 respondents. Off the 72 questionnaires returned only 59 (81.9%) yielded useful data. Only the questionnaire data for horses that had been first evaluated by the Futurity between 2007 and 2013 were retained for analysis.

The sample horse population

Analysis of the data obtained from the questionnaire demonstrated that 59.3% (35/59) of horses evaluated by the Futurity are done so only once, whilst a further 22% are evaluated twice. Only 8.4% and 6.7% are evaluated three or four times respectively, (see Appendix 4a). Horses evaluated more than once equated to 35.5% (21/59) of the sample. Of the horses that were evaluated more than once, 47.6% (10/21) returned to the Futurity evaluations in the subsequent year to their first evaluation, 28.5% (6/21) were evaluated again two years after their first evaluation and 23.8% (5/21) were evaluated three years after their first evaluation (see Appendix 4b and 4c).

The majority of horses evaluated by the Futurity were foals at their first evaluation, which accounted for 45.7% (27/59) of the results. Additionally 16.9% (10/59) were yearlings, 20.3% (12/59) were two year old and 16.9% (10/59) were three year olds. The majority of horses were three year olds at their most recent evaluation and accounted for 71.4% (15/21) of the results (see Appendix 4d).

The majority of premiums awarded at horses first Futurity evaluations were First premiums and account for 38.9% (23/59) of the scores. In horse's most recent evaluations, Higher First premiums were the majority and accounted for 47.6% (10/21) of scores. Analysis of horse scores at their first and most recent Futurity evaluations suggest that 57.1% (12/21) of horse premiums did not change, 28.5% (6/21) were upgraded and 14.2% (3/21) were downgraded. Full horse premium information at Futurity evaluations are summarised in Table 16 and further details are in Appendix 4e.

Table 16: The trends in premium scores awarded at Futurity evaluations.

The trends in premium scores awarded at Futurity evaluations					
First evaluation	Third 3.3%	Second 16.9%	First 38.9%	Higher first 23.7%	Elite 15.2%
Most recent evaluation	Third 4.7%	Second 4.7%	First 28.5%	Higher first 47.6%	Elite 14.2%

The majority of horses at their first evaluation were evaluated for dressage potential and accounted for 42.3% (25/59) of the results. Eventing evaluations accounted for 37.2% (22/59) and 18.6% (11/59) for showjumping (see Appendix 4f). A very small minority of horses changed evaluation discipline between their first and most recent evaluation as follows: one horse entered for eventing in their first evaluation, was entered into dressage in their most recent evaluation and a further two horses both of whom were first entered in the dressage category, were re-evaluated in the eventing category and the showjumping category at their most recent evaluations. No other change of evaluation discipline were reported in the questionnaire data.

Competition data distributions

Analysis from the sample questionnaire data suggests that first premiums were the most common score awarded and accounted for 69.5% (23/59). Elite premium scores were awarded to 15.2% (9/59) of horses and of these horses 7/9 went on to compete in affiliated competitions (Further trends are presented in Table 17).

Table 17: The trends in premium scores of horses which competed in affiliated competitions.

Premium score trends of horses which competed in affiliated competition					
	Third	Second	First	Higher first	Elite
First evaluation	1 out of 2	2 out of 10	16 out of 23	7 out of 14	7 out of 9
	50%	20%	69.5%	50%	77.7%

Analysis of questionnaire data suggests that 64.4% (38/59) of horses competed in unaffiliated competitions and the disciplines they competed in are as follows: 44.7% (17/38) in unaffiliated dressage competitions, 34.2% (13/38) in unaffiliated showjumping competitions and 21.1% (8/38) in unaffiliated eventing competitions. Of the questionnaire total sample 20.3% (12/59) of horses were used for breeding and 8.4% (5/59) did not compete at all (see Appendix 4g).

When considering unaffiliated competition performance only, the majority of Futurity evaluated horses for dressage competed in dressage competitions and accounted for 57.1% (8/14). Of the horses evaluated in the eventing category, 40% (6/15) competed in showjumping competitions and 33.3% (4/15) competed in dressage competitions, only 26.6% (5/15) competed in eventing competitions (see Appendix 4h). However, when considering affiliated competition performance only, the majority of horses competed in the disciplines they were evaluated for; 70.5% (12/17) for dressage, 83.3% (5/6) for showjumping and 46.6% (7/15) for eventing (see Appendix 4i).

When comparing horse premiums with affiliated competition results, this questionnaire suggests that there were no clear associations (see appendix 4j). However, when horse premiums were grouped in two different Chi-Squared tests, there were associations between horse premium awarded and affiliated competition results. In test one, third and second premiums were combined together with first and higher firsts combined together; P=0.023. In test two, third and second premiums were combined together with first, higher first and elite combined; P=0.008, (see Appendix 4k).

When asked if participants horses had reach their desired competition level, 62.2% (33/59) of respondents stated that they had, whereas 37.7% (20/59) stated that they had not (see Appendix 4l). When asked what participants believe prevented their horse from advancing to higher levels, the predominant answers were both at 20% (6/30) was the lack of a suitable rider and the fact that the horse was still progressing through their competition career. Respondents (n=9) indicated that the main reasons that horses did not compete at all was that the horse was being progressed/trained slowly (n=2) or that the horse was sold (n=2) (see Appendix 4m).

Industry opinions on Futurity evaluations

This questionnaire suggests that 52.5% (31/59) of industry opinion states that Futurity premiums are an accurate representation of horse ability, 37.2% (22/59) state that they are not and 10.1% (6/59) are unsure (Appendix 4n). Respondents (n=59) could provide multiple answers when asked what their motivations/reasons were to have their horse Futurity evaluated. Analysis from this questionnaire suggests that the majority of participants used the Futurity evaluations to inform the potential suitability of their horse as a competition horse (35/106) (See Appendix 4o).

Chapter 4. Discussion.

The data distribution

Analysis conducted during this study of Futurity horse evaluation scores and competition performance suggest that only 22% (566/2,559) of Futurity horses evaluated for dressage, eventing or showjumping performance did compete in affiliated competitions. An aim of the Futurity is to identify potential sport horses for competition performance, therefore this large discrepancy between the number of Futurity evaluated horses and affiliated competition performance can suggest that this aim is not being entirely met. However, interestingly the questionnaire results suggested that 55.9% (33/59) of horses did compete in affiliated competitions which is considerably more than suggested from the main population study. This may be a result of two factors, first the questionnaire results do not take into account which discipline horses were evaluated for. Whereas in the main sample, horse competition results were searched for in their respective disciplines that they were evaluated for and therefore records were not used of horses which competed in disciplines other than the one they were evaluated for (these were not searched for due to time constraints limiting access to the data). As a result of these time constraints, it is unknown what proportion may have affiliated in a different discipline other than their evaluation discipline. An additional cause may be contributed to the population which answered the questionnaire. The questionnaire participants can be suggested to have had a greater interest in the Futurity and as such spared time to contribute, whereas the main population study is more representative of the actual population as no sample selection has taken place. Therefore all questionnaire results must be interpreted with caution, however they have been tentatively used to more fully understand the competing trends of Futurity horses.

The majority of the horses evaluated were dressage horses (n=1,049), next were eventers (n=970) and lastly were showjumpers (n=540). These numerical distributions of horse evaluation discipline can be a result of two key factors, namely human preference and horse ability. When considering human preference, sport participation figures can be examined. BD have the most members at 52,000 (British Dressage 2014), BE have 15,000 members (British Eventing 2014), and BS have 16,000 (British Showjumping 2014b). However results from the main data sample suggest that only 62% (556/894) of horses registered with a governing body actually go on to compete. Additionally each governing body offers different priced membership packages and entries which can compete on day ticks, therefore these numbers do not completely represent the number of the competing population. Whilst research by Kearsley (2008b) was undertaken previous to the current membership figures available, it is more comparable across the disciplines as the number of actual competing horses were recorded. Kearsley (2008b) suggests that BD competitions were the most heavily subscribed

per year (n=80,000) compared to BE (n=65,000) and BS (n=50,000). When considering horse ability, it is understood that each discipline places different biomechanical/mental demands on the horse and their physical structure/mental capabilities which make them more suited to a specific discipline (Back and Clayton 2013 and Stewart et al. 2012). Therefore when considering both human preferences and horse ability, it can be suggested that preferences lie mostly with competing horses for dressage competitions and consequently the majority of horses produced may be bred and then evaluated with this discipline and skillset in mind.

Analysis of the Futurity group horses suggest that horses evaluated for dressage competed more than the other disciplines, which is in agreement with Kearsley's (2008b) results which revealed BD competitions had the most entries. Of all Futurity evaluated horses, 24% (257/1,049) of the dressage horses competed, 18% (182/970) of eventers competed and 23% (127,540) of showjumpers competed (all in their respective affiliated disciplines). These results pose the question of why eventers had the largest drop off rate of competing horses at 18%. Results from the questionnaire suggest that in unaffiliated and affiliated competitions, eventers competed in other disciplines more than their own; 73.3% (10/15) and 53.4% (8/15) respectively. Whereas dressage and showjumping horses competed mostly in their respective disciplines; 70.5% (12/17) and 83.3% (5/6) respectively, which may explain why eventers in this study had the largest drop off rate (970 evaluated, 182 with competition records). Similar conclusions were reached by Ray (2012) on the Swedish population of sports horses demonstrated that 85% of eventers also competed in other disciplines. Eventing is a physically demanding sport for both the rider and equid as it consists of three phases requiring multiple all-round skills from the combination (horse and rider) which somewhat explains why eventers may compete in the other disciplines (Kearsley 2008b). Dressage and showjumping competition require very specific (and different) skill sets from the horse and rider combinations, which can explain why these groups kept to their individual disciplines. Furthermore, Ray (2012) also demonstrated that eventing horses started their careers later in life compared to dressage and showjumping horses (typically at 8.5 years of age). Mean age of horses in this study was 7 years old which further explains the large drop off between Futurity evaluated eventers and their limited recorded competition performance. The results of this study clearly demonstrate that different equestrian disciplines have associated with them individual performance trends which can be a result of a combination of the horse's ability, the rider preferences and horse age. However, a key question is why the majority of Futurity evaluated horses did not compete at all and as a result caused this large discrepancy between the number of Futurity horses and affiliated competition performance. Three main causes are discussed.

First it is unclear if all participants who attended the Futurity evaluations intended on affiliating and competing their horses at all. The questionnaire results suggest that the majority of participants did use the Futurity evaluations to gain information on the potential suitability of their horse as a competition prospect, which accounted for 33% (35/106). However interestingly the other answers were not directly linked to competition performance; 24.5% (16/106) used the evaluations to increase the monetary value of their horse, 10.3% (11/106) to inform the potential suitability of their horse as an amateur riding horse, 8.4% (9/106) for vet feedback and 7.5% (8/106) for young horse education. The equestrian trends in GB support the idea that not all participants which attended the Futurity planned on affiliating their horses as a breeding aim in GB is to produce horses for amateur riders due to this being a large proportion of the riding population (Stewart et al. 2012 and BETA 2015). Therefore the discrepancy between Futurity horses and affiliated competition performance can partly be suggested to be a result of equestrian industry trends.

Second, a further cause for the discrepancy between the numbers of Futurity horses with actual competition performance can be partly linked to internal and external influences on the young horse and their environment. The Futurity evaluations are accessible for foals to three year olds. The majority of horses were evaluated as foals (similarly to the questionnaire) and therefore as these young horses progress through their training they can become subject to a number of influencing effects which can affect their competition potential. 37.7% (20/53) of questionnaire participants stated that their horse had not reached their desired competition level and the causes for this were a lack of a suitable rider at 20% (6/30), slow progression at 20% (6/30), horse health/injury at 16% (5/30), horse used for breeding at 16% (5/30) and financial reasons at 16% (5/30). It is not possible to safeguard the horse against all detrimental influences in their training, however the Futurity Equine Bridge programme can assist the highest Futurity scoring horses in reaching their potential by offering financial and training support (see Appendix 1g for further details of the Futurity Equine Bridge).

Third, there is no set age which a horse should begin competing. Mean horse age in this study was 7 years old (ranged from 5 to 12), whereas mean horse age of competing horses ranges from 8.5 to 12 depending on the literature and discipline reviewed (Ray 2012, Stewart 2011, Kearsley 2008a and Ricard and Chanu 2001), therefore as time progresses it can be expected that the number of competition results for the later birth groups will increase with horse age. Figure 4 supports this idea. In 2008/09, the number of Futurity attendees with competition results peaked. However, it is horses born in the 2003/04 birth group which had the greatest number of competition results even though few horses were born in these years. This suggests that with time the 2008/09 birth groups will follow this trend and gain more competition results. These trends were present in all three disciplines (see Appendix 3b). Additionally, the training

and competition trends of riders/owners can further explain the discrepancy between the numbers of Futurity horses with competition results. Media coverage states that as affiliated competitions can be more expensive than unaffiliated competitions, riders are starting their young horses at the lower levels of unaffiliated competition and then waiting until they are mature to affiliate them and progress quickly through affiliated competition, saving money (E-venting 2015). BE statistics by Clissold (2012) supports E-venting (2015) ideas, as results demonstrate that BE80 and BE90 events had fewer starters than BE100, suggesting that competitors progress quickly through the lower levels of the sport. Therefore as horse age, training and competition trends can influence the affiliation of a horse, this knowledge can explain some of the discrepancy between the numbers of Futurity horses with actual competition performances

Futurity premiums

It is clear from analysis of the results of this study that there are many interlinking internal and external factors responsible for the discrepancy between the number of Futurity horses evaluated and the number of Futurity horses with competition results. However the question of greater interest to this study is whether Futurity scores (premiums) are an accurate representation of a horse's potential ability and whether these premiums can be used as a predictor of competition ability. It is important that the BEF Futurity horse premiums are accurate for many reasons; they are used to describe the horse's potential performance ability, they can inform training/breeding decisions and they can increase the horse's monetary value (Kearsley 2008b). All of these reasoning have been demonstrated from the questionnaire to be justifications of why the public attend Futurity evaluations. Of particular current importance is the knowledge that the public are using the Futurity to increase the monetary value of their horses. Currently in GB many horses have little value and are unwanted, affecting equine welfare (World Horse Welfare 2012). Therefore as Futurity premiums have welfare and economic implications, it is paramount that these are representative and reliable. Horse premiums are awarded on a subjective basis, and it is this subjectivity that has been voiced as a limitation to the integrity of the tests used by British horse breeders (Horse Breeders Magazine 2009). Therefore it is this subjectivity which has been investigated to gain a greater understanding of the reliability of Futurity premium scores.

Horses can attended the Futurity more than once and therefore can receive more than one premium score. Of horses evaluated between 2007 and 2010, 20% (2,559/3,177) had been evaluated more than once which may have introduced some bias in the premium scores awarded. Bias may be present by the judges from previous knowledge of the horse, or from altered behaviours from the horse/handlers as a result of experience. Analysis of the

questionnaire data suggests that 57.1% of Futurity premiums do not change from a horse's first evaluation and to their most recent evaluation. As Futurity scores were averaged in cases where a horse had more than one premium score, it can be suggested that this method provides a fuller picture of the horses overall Futurity performance. However, an alternative method of analysis could consist of grouping Futurity horses by age at evaluation, removing attendance bias. Nevertheless to keep to time constraints and for the purpose of simplicity for this preliminary study, horse Futurity scores were average before analysis.

Futurity premiums gained by the horses in this study ranged from 5.6 to 9.9, which is in agreement to Stewart's (2011) Futurity study where a score of 4 was the lowest awarded. First premiums were most commonly awarded at 39.1% (221/566), third premiums at 4.4% (25/566) and 7.1% (40/566) for elite. These results demonstrate that the lower end of the scoring scale is not being utilised and there is a lack of diversity in the premiums awarded, suggesting that either horses are subject to pre-selection based on the Futurity criteria, or as a result of industry pressures on the judges to not excessively penalise a poor performing horse, which if this is the case, has the potential to reduce the reliability of the Futurity evaluations. Futurity factors investigated to determine premium accuracy in the current study were horse sex, age at evaluation, location of evaluation and coat colour.

The majority of horses in this study were males and these accounted for 60.9% (321/566) of horses, which is in agreement with the industry as males (particularly geldings) are preferred in BD, BE and BS competition and a cause of this has been linked to their good temperament and uninterrupted competing performance, as oppose to mares (and some stallions) who are often used for breeding purposes (Viklund et al. 2011 and Kearsley 2008b). There was no relationship between horse sex and the Futurity premium awarded, which differs from results of international and national studies of FPTs. In the RHQT 4yo males scored better in the majority of traits measured (Wallin et al. 2003). Kearsley (2008b) demonstrated that in the YHE stallions scored higher than females (298 horse records) and in the Futurity significant differences were present between the scores awarded to each sex (n=72 horse records). Stallions can be suggested to receive higher scores as they are subject to greater pre-selection; any male showing issues is likely to be gelded. Geldings have therefore been suggested to receive lower scores than stallions in Futurity evaluations (Stewart 2011). A limiting factor in this study was grouping horse sex into two categories (male and female) for consistency of the data and therefore this may explain why no relationship was present between horse premiums and sex. Therefore whilst no correlation was found between horse sex and premiums awarded, it cannot assumed that this not a influencing factor and for that reason must be taken into account when evaluating premium reliability.

Futurity horses are evaluated at different ages and to greater understand the reliability of Futurity premiums, horse age at evaluation must be assessed. In the current study regression analysis established that horse age at evaluation did significantly predict average Futurity score in the showjumping group, however it did not for the dressage and eventing horses. One reason for this discrepancy between the discipline groups can be explained using Stewart's (2011) research suggesting that the youngest evaluated horses received the highest Futurity scores, scores then declined for yearlings and then increase for three year olds. In this study, the highest proportion of showjumping horses were evaluated as foals and less were evaluated as yearlings and two year olds compared to the dressage and eventing group. The showjumping horses additionally were awarded the majority of first premiums in their group suggesting that the higher premiums awarded to foals resulted in the correlation. Previous research about the Futurity suggests that horse age at evaluation significantly effects premiums awarded, particularly for the trait of conformation (Kearsley 2008b and Stewart 2011). In the current study, a further possible reason this discrepancy between the discipline groups could be a result of improvements to the test design, such as more descriptive judging criteria, making the judging process more standardised. Whilst research has demonstrated that some horse premiums can be influenced by horse age, it is unknown which age group receives the most reliable evaluation, as a high premium is not necessarily an accurate one. There are many conflicting findings on the best age to assess young horses. Viklund et al. (2008) and Olsson et al. (2008) suggest testing at 3/4years old to allow tests to be appropriately demanding to assess performance, whilst Schroderus et al. (2006) suggest testing 1/3 year olds as these are less influenced by environment effects and shorten the generation interval. As there is no clear advantage between assessing different age groups of young horses, it can be suggested that having Futurity evaluations open to foals to three year olds can be beneficial to GB as participation is increased whilst the generation interval is decreased.

In Futurity evaluations, the judging panel which attends the evaluations depends on the test location (19 locations). In the current study, evaluation location was not found to predict average Futurity score in any of the discipline groups, which opposes Viklund et al. (2008) results which led to the conclusion that the fixed effect of event location had the greatest influence on FPT results compared to horse sex or age. However, Viklund et al.'s (2008) study was regarding the RHQT and it can be expected that different scoring criteria and judge training schemes are used compared to those in the Futurity evaluations. Interestingly, Kearsley (2008b) demonstrated that YHE mean scores differed significantly across the dates of evaluation. This current study did not account for the date of evaluation and instead accounted for the location, which may explain the opposed results from the YHE. Furthermore

changes have taken place in the Futurity throughout its development, such as Futurity judges undertaking CPD to standardise their skills, the introduction of more descriptive scoring criteria and judges now cannot confer with each other, instead they score individually. These standardisation factors combined may explain why no correlation was found in this study. As evaluation location has not been shown to influence Futurity scores, this can be interpreted as a positive result for the Futurity evaluations as it can suggest that changes in the judging panel do not significantly influence horse scores, however further in-depth study is required to examine this suggestion.

Horse coat colour has been recorded by the Futurity since 2008. In the current study, horse coat colour did significantly predict average Futurity scores in the eventing group, however it was not found to in the dressage and showjumping groups. In the industry, research has demonstrated that competitors have a preference towards the coat colour of the horse, with solid coloured horses being preferred by dressage competitors as it is thought to improve the horses appearance and consequently improving scores (Kearsley 2008b). Findings from more recent research by Fisker-Hansen (2015) is in agreement with Kearsley (2008b) suggesting that the highest Futurity premiums are awarded to solid coloured horses whilst block/spotted horses receive the lowest scores. In the current study the dressage and showjumping group had the highest distribution of coloured/spotted horses and the least solid coloured horses compared to the eventing group, suggesting that the high distribution of solid horses with the least coloured horses may have resulted in the correlation between horse premiums and coat colour in the eventing group only. As research suggests that Futurity premium scores are somewhat influenced by coat colour and that coloured and spotted horses may receive a negative judging bias (Fisker-Hansen 2015), then it can be suggested that the reliability of horse premiums is decreased. To gain a greater understanding of the reliability of horse premiums, this study attempted to determine whether Futurity premiums correlated with horse performance potential and for this to be possible, horse competition performance needed to be objectively measured.

Competition data

The competition data available for the study were obtained from each discipline at each competition level. Analysis of the population sample results suggests that the majority of competition performances were at the lower levels of competition and as the level of competition increased, less horses progressed through the levels (see Figure 4). This is an expected result as BD, BE and BS competition levels are progressive (BD 2015, BE 2015 and BS 2015). Kearsley (2008b) stated that approximately 45% of horses who had competed at BE100 progressed to novice level, whilst in the current study 53% (72/135) of Futurity horses

which competed in BE100 progressed to novice level, demonstrating that Futurity evaluated horses are performing above the general horse population. However in the current study, two dressage grouped horses competed in FEI Intermediate 1 and none progressed higher. In the eventing group only 3 horses competed in BE Advanced competitions and in the showjumping group, 3 competed in BS 1.40m competitions and none progressed higher. Kearsley (2008b) suggests that 45% of intermediate eventing horses progress to advanced competition, whereas only 15% (3/19) did in this study. An aim of the Futurity is to identify potential British horses for elite competition performance, reducing the reliance on international importation of quality horses. Additionally elite scoring horses can enter the Futurity Equine Bridge program which supports horses thought to have the greatest competing potential. Understandably the selection of elite horses must be accurately informed to maximise the efficiency of the Equine Bridge programme. Therefore, whilst results of the study demonstrate that Futurity evaluated horses do compete in elite competitions, the results are concerning as very low numbers of Futurity horses progress to the top levels of competitions. It is important to understand why so few Futurity horses compete in the most elite competitions and a contributing factor may result from the types of equestrians which do/do not participate in the Futurity. Questionnaire results suggests that 37.2% (22/59) of respondents do not believe Futurity premiums are an accurate representation of horse ability, which brings to question whether top breeders/owners/riders have their horses Futurity evaluated. Further study is warranted to increase understanding of what attracts professionals to the Futurity to gain a greater understanding of whether the Futurity is being utilised by top end breeders/owners/riders.

Whilst limited data were available at the highest levels of competition, many records of competition performance were present at the lower and middle ranges. Research has repeatedly demonstrated that correlations are present between the low levels of competition and performance in the higher levels (Stewart et al. 2010, Kearsley 2008a, Hellsten et al. 2006 Aldridge et al. 2001 and Ricard and Chanu 2001). Analysis of results from the lower levels of competition can be suggested to be beneficial to the current study as research has demonstrated that the effect of the rider has less influence compared to advanced levels where greater training and skill are required (Kearsley 2008a). On the other hand, the genetic effect of the horse is less at the lower levels of competition compared to the higher levels (Kearsley 2008a), suggesting study on the records of lower levels of competition performance will not provide as extensive review of horses ability compared to study on records from the more advanced levels. Although as research does suggests that the rider influence is less at the lower levels and as the levels of competition do correlate, the data available in this study were potentially suitable to measure horse performance and have been used to investigate whether Futurity scores were a predictor of performance potential.

Average Futurity scores versus competition data

Analysis of the questionnaire results suggested an association was present between grouped horse premium scores and affiliated competition results. No association was present when premium scores were not grouped, which may be a result of the small sample size (n=58). Whilst these results are a positive finding for the Futurity, they must be interpreted with caution due to the small and potentially selective sample which answered the questionnaire. Analysis of the main sample of data can provide a more comprehensive understanding of Futurity score reliability. In total 74 regression analysis tests were carried out on the main sample data (see Table 5). Horses were grouped by age for analysis to allow each age group to be individually evaluated; a method supported by Kearsley (2008b) who demonstrated that the use of age groups reduces variation in the results of competition scores. In the current study the Futurity average score was used as the predictor and competition performance as the response. In total five of the tests demonstrated significant results.

A key issue with the data in this study was the age of the horses used. Average horse age in this study was 7 years old, whereas the mean horse age at competition ranges from 8.5 to 12 (Ray 2012, Stewart 2011, Kearsley 2008a and Ricard and Chanu 2001). The younger horses used in this study may not have been old enough to have developed their abilities sufficiently to achieve their mature potential and as such limited competition results were available on their records, effecting analysis. Young horses were included in this study as horses which were born in/before 2010 were used to provide a greater sample of horses which had some competition results. Further research is warranted of Futurity scores correlated with competition of horses aged at least 9 years old at competition, compared to 5 years old as used in the current study. However as the Futurity only holds records from 2007 onwards, the database would need to grow before research solely on older horses is possible.

Whilst it is understood that limitations of the data were present, analysis did reveal 5 out of 74 significant correlations between Futurity scores and competition results. These will be discussed in turn.

First significant correlations between Futurity scores and dressage competition results are discussed. Analysis presented that averaged Futurity scores did significantly predict the percentage scores awarded in British Dressage at medium level, as Futurity scores increased by 1point, BD percentage scores at medium level increased by 1.9911 (see figure 5). This result questions why correlations were not present between the lower and higher levels of BD competition results and Futurity scores. It is understood that the genetic effect of the horse is less at the lower levels of competition compared to the higher levels (Kearsley 2008a). This suggests BD records of lower levels of competition performance may not provide as extensive

assessment of horses' ability compared to study on records from the more advanced levels. Additionally as the number of available BD records decreased for horses which reached advanced levels (see Figure 4), nonsignificant data were available to establish significant relationships. A further suggestion of why no additional significant results were established from the dressage group may be a result of the performance measures used in this study (horse points, placings and percentage score). Horse percentage scores demonstrated a significant correlation with dressage scores, no other performance measures did. Horse percentage scores can be considered to be a more accurate representation of the horse's ability compared to points and placings as they measure the performance of the individual horse, whereas the other scores take into the account the performance of other competitors. Further issues were present with the performance data. BD have been recording competition scores of all registered horses since 2011, before this only scores of 60% and above were recorded. The youngest horses in the dressage group were born in 2003, these horses will have reached the age of 8 before all their competition results were recorded. Therefore it can be assumed that some competition records only contained high performance results with some results missing from horse records, creating inconsistencies and bias in the dataset.

A final cause of limited significant correlations between BD scores and Futurity score can be contributed to the subjective nature of dressage competitions. Research by Hawson et al. (2010) investigated Olympic dressage competition scores to understand judging patterns by correlating the collective marks awarded. Results demonstrated considerable variation in the scores, in particular between the scores of horse movement and concluded that judges have considerable difficultly in objectively scoring this subjective discipline. More recent research by Borstal and McGreevy (2014) is in agreement with Hawson et al. (2010) as findings have suggested horses which display incorrect head angles in competition (which is against scoring criteria) achieve the highest scores due to the subjective nature of scoring. Furthermore Kearsley (2008b) demonstrated that the effect of the rider was greatest for the dressage phase of eventing competition which increases the subjectivity of scoring this discipline as the rider can influence the impression of the horse. It is clear that the scoring of dressage competitions is highly subjective, and it is this which may limit correlations between horse FPT scores and competition performance. Results from the RHQT have demonstrated that Swedish young horse scores correlated less with dressage than with showjumping. In 2003, Wallin et al. demonstrated FPT and competition performance correlations of 16/17% for dressage and higher estimations of 23/27% for showjumping. It can be therefore be concluded that dressage performance is highly subjective and as such difficult to predict from limited Futurity records of advanced competing horses.

Second significant correlations between Futurity scores and showjumping competition results are discussed. Interestingly Wallin et al. (2003) demonstrated stronger correlations between FPTs and showjumping competition performance compared with between FPTs and dressage performance. However no significant correlations between Futurity scores and showjumping competition results were demonstrated from this study. There were major issues with the showjumping group performance data. First BS only record results onto their online database of horses which won prize money from their placing or gained a double clear round (no faults gained in their first round and in the jump off), therefore results of poor performing horses are not recorded. This bias in the performance data will limit possible correlations between Futurity scores and competition results. Second, the performance measures recorded by BS are not just a representation of the horse's individual performance, instead factors such as other competitor's performance will influence scoring. Third, as found with results from the dressage group, the genetic effect of the horse is less at the lower levels of competition compared to the higher levels (Kearsley 2008a), suggesting that as BS holds limited records of high levels of competition performance (see Figure 3) this does not provide as extensive review of horses ability compared to study's on more advanced records.

Third significant correlations between Futurity scores and eventing competition results are discussed. Horses from the BE group had the greatest number of correlations between Futurity scores and competition performance, 4 were significant. As Futurity scores increased by 1point, BE penalty point scores significantly decrease/increased (see Table 14). Interestingly the BE group had the greatest number of horses with more advanced competition results (see Figure 3), suggesting that a larger proportion of the dataset had competition results which were more representative of their performance ability, compared to the BD and BS groups. Further strengths were present from the BE data such as BE recording all horse competition results onto their online database since 1998. Additionally It is not surprising that positive correlations were derived from only the performance measure of penalty points as this method measures individual horse performance comprehensively in each phase; whereas the measures of points/placing are influenced by competitor's performance and only record all phases combined. However not all significant correlations between Futurity scores and BE results were expected. As Futurity scores increased by 1point, BE penalty point scores either significantly decrease or increased (see Table 14). The fewer penalty points earnt by a horse, the better they performed in competition, therefore it was unanticipated that penalty points would increase with an increase in Futurity score. BE penalty point scores decreased in the dressage phase of BE90 and novice competition. However, penalty point scores increased in the BE100 dressage phase and the BE90 showjumping phase. These results suggest that whilst Futurity score correlated the most with BE performance, they do not necessarily indicate

that a high Futurity score equals superior competition performance. The attributes of a superior event horse are complex, they must have all-round skill in all three phases of competition and a temperament which is willing to negotiate the demanding event. Therefore it may be that these abilities are difficult to score during a Futurity evaluation and as such contributed to the unexpected results.

It is important to note the R^2 values resulting from correlating Futurity scores with performance measures. R^2 values ranged from 8.55% to 18.41%, suggesting that Futurity score only explained a limited amount of the variability in performance scores. This is understandable as there are many internal external factors which can influence the success of the sports horse, including but not limited to, nutrition, rider, health, time, finance and facilities.

Ultimately there are many variables present in competition, such as the riders influence, rider/horse sex and event location (Kearsley et al. 2008a and Whitaker et al. 2004). As the positive results from this study are so few and sparse they must be interpreted with caution. To summarise, whilst some positive correlations have been demonstrated between Futurity scores and performance measures, further confirmation of the usefulness of using Futurity scores to predict performance in all disciplines is warranted using a larger sample size of older horses with more competition results. A larger sample size will become available as more horses are Futurity evaluated throughout future years. Furthermore, a larger dataset will be available as previously assessed horses have aged and progressed through their competition careers, possibly leading to more competition results becoming available at the higher levels of competition. This is a trend expected to occur as results from the questionnaire suggest that 20% (6/30) of participants are developing their horses slowly, with the aim of affiliating/progressing through the levels once the horse is mature.

Average Futurity component scores versus competition data:

Horse average component scores were used as a further predictor of horse performance at competition. The same performance measures were used as before, see Table 5 (excluding total levels in competition).

Average veterinary score was tested against 74 performance measures. The veterinary mark is awarded to the conformation of the horse's limbs, hooves, muscular-skeletal and other biomechanical factors which can influence the horse's longevity. Results from the RHQT have demonstrated that orthopaedic scores (closely equivalent to the Futurity veterinary score) had the highest correlation with horse longevity compared to other assessed traits (Wallin et al. 2001). Seven correlations between veterinary score and performance were positive in this study (1 in BD, 3 in BS and 3 in BE). However not all correlations were expected as a high Futurity score did not necessarily correlate with superior competition performance (see Table

15). Indicating once again that they may be some limitations in Futurity scoring of desired horse characteristics at evaluations. However as discussed, the trait score of orthopaedic health, which is similar to the veterinary score used in the Futurity, has demonstrated negative correlations with competition results (Wallin et al. 2003). These negative correlations once again imply that as sport horse breeding goals primarily focus on performance success, health and welfare traits may not be prioritised as needed. A limitation of the Futurity veterinary assessment is that the process cannot determine if the horse is prone to heritable diseases/issues, opening it to limitations (Kearsley 2008b and Koenen et al. 2004). Nevertheless questionnaire results demonstrated that 9/27 participants chose veterinary feedback as a reason for using the Futurity evaluations, validating its inclusion in the test. In 2015 the Futurity evaluations underwent changes to the scoring of veterinary mark which is now graded from A (very good) to C (requires substantial management/compromise) and is not included in the overall score. As the veterinary mark informs the health and training requirements of the horse it is an important trait to be included in the Futurity.

Average frame and build scores were tested against 74 performance measures, 5 were positive (3 in BD and 1 in BS and in BE). Frame and build is an assessment of horse conformation which is an important trait of the performance horse as it strongly correlates with performance (Jonsson et al. 2014, Back and Clayton 2013 and Bowing and Ruvinsky 2000). RHQT research demonstrated that horse conformation (n=14,006) scores correlated with horse points and placings in dressage and showjumping (Viklund et al. 2010a). However, in the current study only five positive results were obtained from analysis of frame and build scores versus performance. However, only 46 frame and build scores were available for analysis as this component score was only recorded from 2009 onwards, which is considerably less horses compared to Viklund et al. (2010a). Furthermore, Viklund et al. (2010a) used data of horses born from 1988 to 2007, whereas birth years of 2003 to 2010 were used in this study, limiting its representation of the population.

Average walk, trot and canter scores were tested against 74, 74 and 35 performance measures each. In total 15 of the gait scores correlated with performance (8 in BD, 4 in BE and 3 in BS). Once again not all correlations were expected as a high Futurity score did not necessarily correlate with superior competition performance (see Table 15). As understood, there are issues present in the data sample used in this study and therefore the results cannot be suggested to represent the population. However the distribution trend of the gait scores relates to gait trends shown in other literature regarding the sport horse population. In this study, gait component scores correlated the most with BD performance and the least with BS performance, which is opposite to findings of the RHQT where gait scores correlated the most with showjumping performance (Viklund et al. 2010a and Wallin et al. 2003) and the least with

eventing performance (Ray 2012). FPT gait scores should highly correlate with dressage performance (as it is these which are judged). Equally however the gaits of the showjumping horse are important as FPT scores of walk and trot have been demonstrated to positively correlate with canter score (Hellsten et al. 2006) and canter ability correlates with jumping ability (Back and Clayton 2013 and Hellsten et al. 2006).

Average jump score was tested against 26 performance measures of which only one was positive (average BE100 penalty points in the showjumping phase). As eventing competition consists of two jumping phases (cross country and showjumping), it is understandable that the jump trait correlated with this discipline. No correlations were demonstrated with BS results, however only 13 horses had the jump component score. In RHQT studies, jump traits have highly correlated with showjumping performance (Viklund et al. 2010a and Wallin et al. 2003), and moderately correlated with eventing competition (Ray 2012). Interestedly, research by Kearsley (2008b) demonstrated that the jumping trait score is not significantly affected by the date of Futurity evaluation, whereas other component traits were, suggesting it is easier to objectively judge. As the jumping trait can be objectively evaluated and has demonstrated high correlations with performance in RHQT studies (compared to gait traits), once again it can be suggested that further research is needed with a greater sample size to present a clearer and more reliable picture of the results obtained from this study

The final component trait athleticism was tested against 74 performance measures, 7 were positive (5 in BD and 2 in BE). Only one of these correlations between Futurity scores and competition results were unexpected, this was in the BD dataset (see Table 15). The athleticism trait had the most positive results from a single trait which is a promising result as in 2014 the trait was scored out of 20 (as oppose to 10) and therefore contributing more to the overall Futurity score. The Athleticism trait has additionally been demonstrated to correlate with all other component scores, excluding veterinary (Stewart 2011 and Kearsley 2008b). The phenotypic correlations suggest two effects; either that the genes responsible for the trait are linked to the genes which are responsible for other the component traits or that if a horse is judged as athletic in the futurity, it will then be scored as a good horse; it may be a combination of both. Nevertheless selection based on athleticism does suggest that other skills of the sports horse (suitable conformation and paces) will also be selected. Research supports the double weighting of athleticism scoring in the Futurity, as do the results of this study since the trait expectedly correlated with the most performance measures.

The trait type and temperament was removed from the Futurity in 2009, although type is still assessed under the other traits, but temperament was deemed too subjective to assess on the day of evaluation. The removal of this trait is questionable as research has demonstrated

that it strongly correlates with the athleticism score (Stewart 2011) and is very important to the owner/rider as it is associated with performance (Randle et al. 2015 Kearsley 2008b, Stewart et al. 2012 and Koenen et al. 2004). In a questionnaire distributed by Kearsley (2008b), BD reported temperament was an important trait of the dressage horses, BE and BS stated it wasn't as important. Interestedly results of the current study demonstrated dressage performance had the greatest number of unexpected correlations between performance measures and average Futurity component score (5/17). As temperament is a key dressage trait, the unexpected correlations may result from the Futurity scores not taking into account horse temperament and as such did not provide a full analysis of the horse. It can be suggested that the temperament trait is reintroduced to the Futurity criteria to provide an extensive analysis of the horse, however as it is often described as suitable or good (Stewart et al. 2010), more objective terminology is needed with reintroduction. Research by Olsson (2010) investigated different methods of objectively evaluating horse temperament, suggesting tests which measure reactivity to stimuli are easier to objectively assess than those which measure a horse's cooperation with a human, as the human influences the results.

Ultimately results from this study have demonstrated interesting findings. First it has become apparent that with growth the Futurity dataset will develop into a more comprehensive data source with the addition of more evaluated horses with more competition results at advanced levels of competition. Second performance measures usefulness has been evaluated. In eventing all significant positive results were demonstrated from penalty points which is not surprising as this method measures individual horse performance; it is not influenced by other competitors. Kearsley (2008a) first introduced the method of analysing eventing horses from penalty points, demonstrating their usefulness in predicting performance heritabilities in all phases of competition. Horse percentage scores in dressage demonstrated the only other significant correlation between average Futurity score and performance. Percentage scores are awarded based on the individual and not influenced by other competitors, it is this evaluation method that is recommended by Stewart et al. (2010). However unlike penalty point scores in BE, percentage scores are awarded in a much more subjective manner due to the nature of dressage scoring. BS had no performance measures which evaluated individual performance, contributing to why limited correlations between scores were found. A faultless scenario for this study would include a lager Futurity dataset with horses over the age range of 8.5 to 12 (average competition age) where governing bodies record all competition result of all horses and all use a performance measure which is based on individual performance. In time this scenario may become accessible with growth of the Futurity and an increase in competition scoring. However as of current BE and BD are the only governing bodies which contain a performance measure evaluating induvial performance only, BS do not.

Chapter 5. Conclusion.

Summary

Analysis of the results of this study suggest Futurity evaluations are attracting participants who are keen understand their horse's competition potential, veterinary health and monetary value, amongst other reasons (see Appendix 4o). As participants are attending Futurity evaluations to receive professional feedback on their horse, it is important that this information is accurate as it has welfare and economic consequences. Analysis of Futurity premium scores suggest that they may be influenced by horse related variables such as horse age and coat colour. However interestingly, evaluation location (used to describe the judging panel) was not suggested to influence scores. Previously the effect of Futurity evaluation date significantly influenced premium scores (Kearsley 2008b), however since this the Futurity tests have undergone developments to reduce test subjectivity; explaining why evaluation location may not have influenced Futurity scores in this study (BEF 2015b and Horse Breeders Magazine 2009). In 2015 further measures have been introduced to control for subjectivity (after this study was conducted). Linear scoring has been introduced and is said to promote breeding and performance success from descriptive and objective scoring criteria leading to a greater understanding of the horse's phenotypic traits. Research by Stock (2013) suggests that linear scoring is promoting genetic gain of German sport horses as the characterising of phenotypes has improved from phenotypic analysis. Stock (2013) additionally suggests that as more studbooks are incorporating linear scoring into their systems, transparency and collaboration between organisations is improved, allowing greater genetic gain of the sport horse to become possible. Over a decade ago Koenan et al. (2004) stated that the varying subjective scoring of sport horses was a limitation of the industry, and that standardisation would lead to more informed breeding and training practices. Therefore, the introduction of linear scoring in the Futurity is likely to be a promising step for British Breeding and future success.

The main aim of the study was to establish if Futurity scores correlated with competition performance to gain a greater understanding of the tests usefulness. Limited data were available due to a large proportion of Futurity horses not having competition results (78%). The reasons for this have been contributed to human competing preferences, young horses still developing and varying horse abilities. The majority of competing horses did so in dressage competition and the minority in showjumping competitions. Limited positive correlations were found between average Futurity scores and component scores with competition performance; the majority of positive correlations found in the BE group.

This study highlighted a number of important trends for the industry. First, it is now understood that with time the Futurity will develop into a more comprehensive data source. Second,

performance measures usefulness has been evaluated and it is understood that the most appropriate method to measure horse performance is a system which evaluates the individual's performance and does not take into account the ability of other competitors, such as penalty points and percentage scores as used by BE and BD respectively. Last, as of current it has been not been establish if a high Futurity score equals a high competition score as a proportion of the results of this study demonstrated the opposite, highlighting the difficulty of evaluating a young sport horse in a brief Futurity assessment. To thoroughly understand the potential of Futurity evaluations more comprehensive datasets are needed such as a larger Futurity dataset with horses over the age range of 8.5 to 12 (average competition age). Governing bodies which record all horse competition results all use performance measures based on individual performance only. In time this may become accessible as sport horse breeding in Britain develops.

In Sweden, the RHQT is extensively used to successfully evaluate the phenotypic potential of their sport horse population and research has confidently confirmed the usefulness of their evaluations. In the United Kingdom, the BEF hope to incorporate the Futurity evaluations to form phenotypic evaluations of the British sport horse. Whilst the results of the current research cannot suggest that Futurity scores are a predictor of competition potential from the results established, further research is warranted.

Conclusion

Few horses with Futurity evaluations records are competing in affiliated competitions. Those horses which do affiliate mostly compete in the low to middle ranges of competition, potentially as a result of the mean age of horses in this study being only 7 years old, meaning they have not currently progressed to their potential. The young horses used in this study is thought to be a main contributing factor of the limited positive correlations demonstrated between Futurity scores and measures of performance. However what this study has identified is the importance of governing bodies to comprehensively and consistently record horse performance at competition and to use a transparent recording system.

Chapter 7. References.

Aceto, H., and Parente, E. (2012) “Using Quarterly Earnings to Assess Racing Performance in 70 Thoroughbreds after Modified Laryngoplasty for Treatment of Recurrent Laryngeal Neuropathy”, *Veterinary Surgery*, 41(6), 689-695.

Albertsdottir, E., Eriksson, S., Nasholm, A., Strandberg, E., and Arnason, T. (2008) “Genetic analysis of competition data on Icelandic horses, *Livestock Science”*, 114, 181-187.

Aldridge, L.I., Kelleher, D.L., Reilly, M., and Brophy, P. (2001) “Estimation of the genetic correlation between performances at different levels of show jumping competitions in Ireland”, *Journal of Animal Breeding and Genetics*, 117(1), 65-72.

Back, W., and Clayton, H.M. (2013) “Equine Locomotion”, 2ed. Saunders Ltd: China.

Berrey, E., Desliens, F., Poirel, D., Biau, S., Lemaire, S., Rivero, J., and Langlois, B. (2002) “Early evaluation of dressage ability in different breeds”, *Equine Veterinary Journal*, 34, 319-324.

Becker, A.C., Stock, K.F. and Distl, O. (2011) “Genetic correlations between free movement and movement under rider in performance tests of German Warmblood horses”, “Livestock Science”, 142, 245-252.

Bowing, A.T., and Ruvinsky, A. (2000) “The genetics of the horse”. CAB: Oxon.

Braam, A., Nasholm, A., Roepstorff, L., and Philipsson, J. (2011). “Genetic variation in durability of Swedish Warmblood horses using competition results”, *Livestock Science*, 142(1), 181-187.

Braam, A. (2011) “Breeding for durable riding horses using competition statistics”, *Swedish University of Agricultural Sciences*, 147, 1-32.

British Breeding: British Equestrian Federation Futurity Results database, (2015), [online], http://www.britishbreeding.org/futurityResults.aspx, data accessed: 08.10.15.

British Dressage. Members Handbook, (2015). [Online], http://www.britishdressage.co.uk/uploads/File/Rules%202015/BD%20Memebers%20Handbook%202015%20Web%20File.pdf, date accessed: 03.12.15.

British Dressage: Results and Rankings, (2014), [online], http://www.britishdressage.co.uk/results/latest_results, date accessed: 28.10.14.

British Equestrian Federation Futurity: Factsheet 5, (2015b), [online], http://www.bef.co.uk/repository/downloads/Horses/Futurity/2015/Futurity_Factsheet_5_FV.pdf, date accessed 20.10.15.

British Equestrian Federation Futurity: Futurity Equine Bridge (2014c), [online], http://www.bef.co.uk/repository/downloads/Horses/Bridge/Futurity_Equine_Bridge_2014__FV.pdf, date accessed: 20.10.15.

British Equestrian Federation: BEF Futurity, Factsheet 1, (2015a), [online], http://www.bef.co.uk/repository/downloads/Horses/Futurity/2015/Factsheet_1.pdf, date accessed: 22.10.15.

British Equestrian Federation: BEF Futurity, Factsheet 4, (2014b), [online], http://www.bef.co.uk/repository/downloads/Horses/Futurity/2014/Factsheet_4_Eventing_2014.pdf, date accessed: 28.12.14.

British Equestrian Federation: BEF Futurity, Factsheet 4, (2015d), [online], http://www.bef.co.uk/repository/downloads/Horses/Futurity/2015/Factsheet_4.pdf, date accessed 22.10.15.

British Equestrian Federation: BEF Futurity, Factsheet 8, (2014a), [online], http://www.bef.co.uk/repository/downloads/Horses/Futurity/2014/Factsheet_8_2014.pdf, date accessed 09.10.14.

British Equestrian Federation: BEF Futurity, Factsheet 8, (2015c), [online], http://www.bef.co.uk/repository/Factsheet_8.pdf, date accessed 22.10.15.

British Equestrian Trade Association: National Equestrian Survey 2015 shows increased consumer spending, (2015), [Online], http://beta-uk.org/pages/news-amp-events/news/national-equestrian-survey-2015-shows-increased-consumer-spending.php, date accessed: 14.10.15.

British Eventing, Horse search, (2015a). [Online], http://www.britisheventing.com/search.asp?section=156&sectionTitle=Search, date accessed: 27.8.15.

British Eventing. Member Handbook, (2015b). [Online], http://www.britisheventing.com/documents.asp?section=145&sectionTitle=Download+Area#category0001200000010002, date accessed: 03.12.15.

British Eventing: Rules and Scoring, (2009), [online], http://www.britisheventing.com/asp-net/page.aspx?section=681&itemTitle=The+Rules+of+Each+Event+Phase+and+How+They+are+Scored, date accessed: 26.10.15.

British Eventing: The Olympic Eventing format for London 2012, (2012), [Online], http://www.britisheventing.com/asp-net/page.aspx?section=1122, date accessed: 22.09.15.

British Eventing: Directors statement, report and financial statements. (2014b), [online], file:///C:/Users/Lauren/Downloads/Complete_Directors_Report_Statement_2014_.pdf, date accessed: 24.02.16.

British Showjumping, Horse search, (2015a) [online], https://members.britishshowjumping.co.uk/horses/search, date accessed: 27.08.15.

British Showjumping, Points allocations and point's limits – update, (2014a). [online], http://www.britishshowjumping.co.uk/media/news/points-allocations-and-points-limits---update, date accessed: 27.08.15

British Showjumping. Member Handbook. (2015b), [Online], http://www.britishshowjumping.co.uk/_files/bs%20rule%20book_sm.pdf, date accessed: 03.12.15.

Buxadera, A., and Mota, M.S. (2008) "Variance component estimations for race performance of thoroughbred horses in Brazil by random regression model", *Livestock Science*, 117(2), 298-307.

Carta, A., Casu, S., and Salaris, S. (2009) “Current state of genetic improvement in dairy sheep”, *Journal of Diary Science*, 92, 5814-5833

Clausen, A., and Whitaker, T. (2009) “Analysis of the trait scoring data in Futurity Evaluations (2005-2008) for young sports horses” Writtle College Bachelor thesis.

Clissold, J. (2012) “Safety review and statistics”, [online], https://www.britisheventing.com/library-media/documents/safetyreport.pdf, date accessed; 24.02.16

Coffey, M., Kryzelewski, T., Roughsedge, T., and Mrode, R. (2007) “EGENES and BASCO Nation Beef Evaluations in the UK”, *Scotland Rural College*, 1-6.

Crossman, G.K., (2010) “The organisational landscape of the English horse industry; a contrast with Sweden and the Netherlands”, University of Exeter PhD, [online], https://ore.exeter.ac.uk/repository/bitstream/handle/10036/111475/CrossmanG.pdf, date accessed: 20.10.15.

Dubois, C., and Ricard, A. (2007) “Efficiency of past selection of the French Sport Horse: Selle Français breed and suggestions for the future”, *Livestock Science*, 112(1), 161-171.

Ducro, B.J., Koenen, E.P.C., Tartwijk, J., and Arendonk, J. (2007) “Genetic relations of First Stallion Inspection traits with dressage and show-jumping performance in competition of Dutch Warmblood horses”, *Livestock Science*, 107(1), 81-85.

Ducro, B.J., Speelman, H., Koenen, E.P.C., and Van Tartwijk, J.M.F.M. (2002) “Genetic relationship between the first stallion inspection and the dressage and show jumping competition results in the Dutch Warmblood riding horse population”, *Livestock Production Science,* 30, 379-382.

Epstein H., and Mason I.L. (1984), “Cattle. In: Evolution of Domesticated Animals”, Longman Inc: New York.

E-Venting. Has BE90 gone mad? (2015). [online], http://e-venting.co.uk/2015/06/has-be90-gone-mad/, date accessed: 24.02.16

EquiMed. (2010) “Equine reproductive maturity in mares and stallions”, [online], http://equimed.com/health-centers/reproductive-care/articles/equine-reproductive-maturity-in-mares-and-stallions , date accessed: 19/10/2016

Falconer, D., and Mackay, T. (1995) “Introduction to Quantitative Genetics”, 4th ed, Longman: Essex.

Fisker-Hansen, A.., Randle, H., and Dixon, J.L. (2015) “An investigation of equine coat colour bias in assessment of potential performance horses”, paper presented at *The International Society for Equitation Science,* Vancouver, 5-8 August.

Gelinder, A., Skoglund, A., Nasholm, A., and Phillipsson, J. (2001) “Relationships bewteen stallion performance test results and sport results in dressage and show jumping”, paper presented at *Europe Federation of Animal Science*, Hungary, 26-29 August.

Gelinder, A., Thoren, E., Nasholm, A., and Phillipson, J. (2002) “Relationship between test results at 3 and 4 years of age for Swedish riding horses”, *Livestock Production Science*, 30, 375-378.

Gjedrem, T., and Baranski (2009). “Selective Breeding in Aquaculture: an Introduction”, Springer: London.

Goddard, M., Hayes, B., and Meuwissen, T. (2010) “Genomic selection in livestock populations”, *Genetics research*, 5(6), 413-421.

Guide Dogs: Selection and monitoring of Breeding Stock. (2015), [Online], http://www.guidedogs.org.uk/aboutus/national-breeding-centre/nbc-revealed/selection-and-monitoring-of-breeding-stock/#.VedOZvlViko, accessed on; 5.9.15.

Hamann H., and Distl, O. (2007) “Genetic variability in Hanoverian warmblood horses using pedigree analysis”, *Journal of Animal Science*, 86(7), 1503-1513.

Hawson, L., McLean, A., and McGreevy, P. (2010) “Variability of scores in the 2008 Olympic dressage competition and implications for horse training and welfare”, *Livestock Science,* 5(4), 170-176.

Hayes, B., Harris, A., and Goddard, M. (2013) “The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaption”, *Trends in Genetics*, 29(4), 206-214.

Hellsten, E.T., Viklund, A., Koenen, E.P.C., Ricard, A., Bruns, E., and Philipsson, J. (2006) “Review of genetic parameters estimated at stallion and young horse performance tests and their correlations with later results in dressage and show jumping competition”, *Livestock Science*, 103, 1-12.

Holmstrom, M., and Philipsson, J. (1993) “Relationships between conformation, performance and health in 4 year old Swedish Warmblood Riding Horses”, *Livestock Production Science*, 33, 293-312.

Horse Breeders Magazine: The BEF Futurity Evaluations: The Future for British Breeding with Jan Rogers & Celia Clarke, (2009), [online], http://www.horsebreedersmagazine.com/#/futurity-interview/4537595051: date accessed: 07.10.14.

Iwona, J., Wilk, I., Zalewska, E., and Bocian, K. (2015) “Correlations between the behaviour of recreational horses, the physiological parameters and summer atmospheric conditions” *Animal Science Journal*, 86, 721-728.

Jonsson, L. (2013) “Orthopaedic Health, Conformation and Longevity in Riding Horses”, Swedish University of Agricultural Sciences Doctoral Thesis.

Jonsson, L., Nasholm, A., Roepstorff, L., Egenvall, A., Dalin, G., and Philipsson, J. (2014) “Conformation traits and their genetic and phenotypic associations with health status in young Swedish warmblood riding horses”, *Livestock Science*, 163, 12-25.

Jorgensen, G.H., and Boe, K. (2007) “A note on the effect of daily exercise and paddock size on the behaviour of domestic horses (Equus caballus)”, *Applied Animal Behaviour Science*, 1(2), 166-173.

Kearsley, C., Woolliams, J.A., Coffey, M.P., and Brotherstone, S. (2008a) “Use of competition data for genetic evaluations of eventing horses in Britain: Analysis of the dressage, show jumping and cross country phases of eventing competition”, *Livestock Science*, 118, 72-81.

Kearsley, C. (2008b) “Genetic Evaluation of Sport Horses in Britain”, University of Edinburgh PhD, [online] https://www.era.lib.ed.ac.uk/bitstream/1842/3154/1/KearsleyC%20PhD%20thesis%2008.pdf, date accessed: 09.11.14.

Koenen, E.P.C., Aldridge, L., and Philipsson. (2004) “An overview of breeding objectives for Warmblood sport horses”, *Livestock Production*, 88, 77-84.

Langlois, B., and Blouin, C. (2004), “Practical efficiency of breeding value estimations based on annual earnings of horses for jumping, trotting, and galloping races in France”, *Livestock Production Science*, 87(2), 99-107.

Luehrs-Behnke, H., and Kalm, E. (2002a) “Genetic associations among traits of the new integrated breeding evaluation method used for selection of German warmblood horses”, *Veterinarija ir Zootechnika*, 18, 90-93.

Luehrs-Behnke, H., Rohe, R., and Kalm, E. (2002b) “Estimation of genetic parameters for traits used in the integrated breeding evaluation of Germen Warmblood horses”, paper presented at *Europe Federation of Animal Science*, Egypt, 1-4 September.

Military Police: Military Working Dog Program. (1993), **3**0 September, 2-3.
Molina, A., Valera, M., Santos, R., and Rodero, A. (1999) “Rodero Genetic parameters of morphofunctional traits in Andalusian horse” *Livestock Production* Science, 60, 295–303.

Munsters, C., Iwaarden, A., Weeren, R., and Marianne, M. (2014) “Exercise testing in Warmblood sport horses under field conditions”, *The Veterinary Journal*, 1, 11-19.

Negussie, E., Koivula, M., and Mantysaari, E. (2006) “Genetic parameters and single versus multi-trait evaluation of udder health traits”, *Animal Science*, 56(2), 73-82.

Neyround, E., and Daniels, S. (2013) “Does Body condition and muscle condition score correlate to British Equestrian Federation Futurity evaluation score”, paper presented at Alltech-Hartpury Conference 2013.

Olsson, E., Nasholm, A., Strandberg, E., and Philipsson, J. (2008) “Use of field records and competition results in genetic evaluation of station performance tested Swedish Warmblood stallions”, *Livestock Science*, 117, 287–297.

Olsson, E.G., Arnason, T., Nasholm, A., and Philipsson, J. (2000) “Genetic parameters for traits at performance test of stallions and correlations with traits at progeny tests in Swedish warmblood horses”, *Livestock production science*, 65, 81-89.

Olsson, K. (2010) “A review of methods used to measure temperamental characteristics in horses”, *Swedish University of Agricultural Sciences*.

Posta, J., Malovhr, S., Mihok, S., and Komlosi, L. (2010) “Genetic variation in durability of Swedish Warmblood horses using competition results”, *Journal of Animal Breeding and Genetics*”, 127(4), 280-288.

Pryce, J., Hayes, B., and Goddard, M. (2011), “Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information”, *Journal of Dairy Science*, 95(1), 377-388.

Ray, B.A. (2012) “Genetic analysis of eventing data in the in Swedish Warmblood population”, Swedish University of Agricultural Sciences Master’s Thesis.

Ricard, A., and Chanu, I. (2001), “Genetic parameters of eventing horse competition in France”, *Genetics Selection Evolution*, 33, 175-190.

Roos, L., Hinrichs, D., Nissen, T., and Krieter, J. (2015) “Investigations into genetic variability in Holstein horse breed using pedigree data”, *Livestock Science*, 177, 25-32.

Rothmann, J., Christensen, O.F., Sondergard, E., and Ladewig, J. (2014) “A note on the heritability of reactivity assessed field tests for Danish Warmblood horses”. *Journal of Equine Veterinary Science*, 34, 341-343.

Safari, E., Fogarty, N.M., Gilmour, A.R., Atkins, K.D., Mortimer, S.I., Swan, A.A., Brien, F.D., Greeff, J.C., and Van Der Werf, J.H.J. (2007) “Genetic correlations among and between wool, growth and reproduction traits in Merino sheep”, *Journal of Animal Breeding and Genetics*, 124(2), 65-72.

Schroderus, E. (2006) “Estimates of genetic parameters for Finnhorse and Standardbred trotter foals”, University of Helsinki Master’s Thesis.

Sinn, D., Goslling, S., and Hilliard, S. (2010) “Personality and performance in military working dogs: Reliability and predictive validity of behavioural tests”, *Applied Animal Behaviour Science*, 127, 51-65.

Stewart, I. (2011) “Genetic Evaluation of Traits Recorded in Sport Horses in GB”, University of Edinburgh PhD, [online], https://www.era.lib.ed.ac.uk/bitstream/1842/7971/2/Stewart2012.pdf, date accessed: 10.11.14.

Stewart, I., White, M., Gilmour, A., Thompson, R., Woolliams, J., and Brotherstone, S. (2012) “Estimating variance components and predicting breeding values for eventing disciplines and grades in sport horses”, *Animal*, 6(9), 1377-1388.

Stewart, I., Woolliams, J.A., and Brotherstone, S. (2010) “Genetic evaluation of horses for performance in dressage competitions in Great Britain”, *Livestock Science*, 128, 36-45.

Stock, K., and Distl, O. (2006) “Genetic correlations between conformation traits and radiographic findings in the limbs of German riding horses”, *Genetics Selection* Evolution, 38, 657-671.

Stock, K. (2013) “Linear profiling in the warmblood horse – review and preview” *WBFSH seminar*, October 7th.

Suffolk Horse Society: The Suffolk Horse. N.d. [Online], http://suffolkhorsesociety.org.uk/management/, accessed on; 5.9.15.

Sun, C., Madsen, P., Lund, M.S., Zhang, Y., Nielsen, U.S., and Su, G. (2010) “Improvement in genetic evaluation of female fertility in dairy cattle using multiple-trait models including milk production traits”, *Journal of Animal Science*, 88(3), 871-878.

Suontama, M., Saastamoinen, M., and Ojala, M. (2009) “Estimates of non-genetic effects and genetic parameters for body measures and subjectively scored traits in Finnhorse trotters”, *Livestock Science*, 1(3), 205-209.

Suzuki, D., and Griffiths, A. (2000) “Introduction to Genetic Analysis 7th ed”, Oxford: Freeman and Company.

Taberlet, P., Valentini, A., Rezaei, H., Naderi, S., Pompanon, F., Negrini, R., and Marsan, P. (2008), “Are cattle, sheep, and goats endangered species?” *Molecular Ecology*, 17(1), 275-284.

Tavernier, A. (1992) “Is the performance at 4 years in jumping informative for later results”, paper presented at *Europe Federation of Animal Science,* Spain, 13-17 September.

The British Percheron Society: Characteristics of the British Percheron. 2009, [Online], http://www.percheron.org.uk/about/percheron.html, accessed on; 5.09.15.

Thiruvenkadan, A.k., Kandasamy, N., and Panneerselvam, S. (2009) “Inheritance of racing performance of Thoroughbred horses”, *Livestock Science*, 121(2), 308-326.

Trybocka, R. (2010) “Character assessment testing to test suitability for guide dogs”, *Veterinary Nursing Journal*, 25(11), 32-33.

Tsuruta S., Aguilar I., Misztal I., Lawlor T. J. (2011), “Multiple-trait genomic evaluation of linear type traits using genomic and phenotypic data in US Holsteins”, *Journal of Dairy* Science, 94,4198–4204.

Velie, B., Hamiton, N., and Wade, C. (2015) “Heritability of racing durability traits in the Australian and Hong Kong Thoroughbred racing populations”, *Equine Veterinary Journal,* 82, 61-71.

Viklund, A., Braam, A., Nasholm, A., Strandberg, E., and Philipsson, J. (2010a), “Genetic variation in competition traits at different ages and time periods and correlation with traits at field test of 4-year-old Swedish Warmblood horses”, *Animal*, 4(5), 682-691.

Viklund, A., Hellsten, E.T., Nasholm, A., Strandberg, E., and Philipsson, J. (2008) “Genetic parameters for traits evaluated at field tests of 3- and 4-year-old Swedish Warmblood horses”, *Animal*, 2(12), 1832-1841.

Viklund, A., Nasholm, A., Strandberg, E., and Philipsson, J. (2010b) “Effects of long-time series of data on genetic evaluations for performance of Swedish Warmblood riding horses”, *Animal*, 4(11), 1823-1831.

Viklund, A., Nasholm, A., Strandberg, E., and Phillipsson, J. (2011) “Genetic trends for performance of Swedish Warmblood horses”, *Livestock science*, 141(2), 113-122.

Villemereuil, P. (2012) “Estimation of a biological trait heritability using the animal model”, *Animal Ecology,* 79, 1-35.

Visser, E.K., Reenen, C.G., Rundgren, M., Zetterqvist, M., Morgan, K. and Blokhuis, H. (2003) “Responses of horses in behavioural tests correlated with temperament assessed by riders*” Equine Veterinary Journal,* 35(2), 176-183.

Wallin, L., Strandberg, E., and Phillipsson, J. (2001) “Phenotypic relationship between test results of Swedish Warmblood horses as 4-year-olds and longevity”, *Livestock Science*, 68(2), 97-105.

Wallin ,L., Stranberg, E.,and Phillipsson, L. (2003) “Genetic correlations between field test results of Swedish Warmblood Riding Horses as 4-year-old and lifetime performance results in dressage and show jumping” *Livestock Production,* 82, 61-71.

Weller, J., Stoop, W., Eding, H., Schrooten, C., and Ezra, E. (2015) “Genomic evaluation of a relatively small dairy cattle population by combination with a larger population”, *Journal of Dairy Science*, 98(7), 4945-4955.

Whitaker, T., Hill, J., and Shearman, J. (2004) “Scoring analysis of completing pre-novice event horses at six selected events”, *Exercise and Comparative Exercise Physiology*, 1(3), 185-192.

Whitaker, T., Olusola, O., and Redwin, L. (2008) “The influence of horse gender on eventing competition performance”, *Comparative Exercise Physiology*, 5(2), 67-72.

World Horse Welfare: Left on the verge, the approaching equine crisis in England and Wales (2012), [Online], https://www.horseworld.org.uk/uploads/documents/1350371514_Horsecrisisreportfinal.pdf, date accessed; 04.04.16

Randle, H. 2015. Personality and performance: the influence of behavior**,** in Williams, J and Evans, D. (2015) “Training for equestrian performance”, Wageningen Academic Publishers:Wageningen.

World Breeding Federation for Sports Horses: Dressage horse and studbook rankings, (2015a), [Online], http://www.wbfsh.org/files/Dressage_Studbook_September_final_2015.pdf, date accessed: 20.11.14.

World Breeding Federation for Sports Horses: Eventing horse and studbook rankings, (2015b), [Online], http://www.wbfsh.org/files/Eventing_Studbook_September_final_2015.pdf, Date accessed: 20.11.14.

World Breeding Federation for Sports Horses: Jumping horse and studbook rankings, (2015c), [Online], http://www.wbfsh.org/files/Jumping_Studbook_September_final_2015.pdf, date accessed: 20.11.14.

World Breeding Federation for Sports Horses: Horse and studbook rankings, (2009), [Online], http://www.wbfsh.org/GB/Rankings/Horse%20and%20studbook%20rankings/2009.aspx, date accessed: 01.11.14.

World Breeding Federation for Sports Horses: Horse and studbook rankings, (2005), [Online], http://www.wbfsh.org/GB/Rankings/Horse%20and%20studbook%20rankings/2005.aspx, date accessed: 01.11.14.

Appendices.

Section 1. Literature review

1a. Warmblood breeding objectives

Warmblood Breeding Objectives of European Countries	
Baden-Württemberg	The breeding objective includes two traits. The first is called “sport”. The aim is to meet the expectations of ambitious competition-riders. There is a specialisation either for dressage or for show jumping. The second trait is called “nice and easy”. The aim is to meet the requirements of the pleasure-rider. Selection criteria are beauty and easy handling.
Bavarian Warmblood	Riding horse that performs well in dressage, show jumping and eventing with an elegant conformation and willingness to work, reliable temperament, a good health, high durability and good fertility.
Belgian Warmblood	A noble modern and correctly-built warmblood horse with a rectangular frame, big outlines and good basic paces. The horse should be pleasant to ride and have an unobjectionable character, so that it can be used by any rider, both as pleasure horse and as a performance horse. There has to be a balance between conformation, performance and health.
Danish Warmblood	A riding horse with a big performance ability in either dressage or show jumping, able to compete at international level. The horse is noble, sizeable, subtle with good health in and high reproductive ability.
Finnish Warmblood	A horse that is suitable for use in dressage, show jumping and eventing at national and international level on basis of its type, conformation, gaits, character and soundness
Hanoverian	A rideable, noble, big framed and correct warmblood horse that, based on his qualities, temperament and character is especially suitable to use as a performance and leisure horse. The breeding of sport horses with a good ability for the disciplines dressage, jumping, eventing and driving is encouraged.
Holsteiner	An athletic long-shaped and expressive horse with the ability for sport performance at national and international level, especially show jumping. Desired is a powerful elastic and good jumping, showing good control and intelligence. The movements must show the typical knee action of the Holstein horse.
Hungarian Sporthorse	A noble, attractive, strong marketable horse for use under saddle and driving suitable for all the different branches of equestrian sports approved by the FEI, i.e. dressage, show jumping and eventing both at national and international competitions.
Irish Sport Horse	Sound and athletic horses that are capable of competing in show jumping, eventing and dressage at both national and international level.

Dutch Warmblood	Sport horses performing at the highest level of international competition (dressage and/or jumping) with a good constitution and a functional and well-shaped conformation.
Dutch Riding Horse Studbook	Riding horses with (anglo) arabic blood that contributes to sport and able to perform in competitions.
Norwegian Warmblood	A quality sport horse of good type which is correct, has a good frame and is sound. The horse should have a good temperament, a good attitude to discipline, co-operative and free forward movement, suitable for performance in dressage, show jumping and eventing. Conformation, performance capability, soundness, temperament and fertility are all to be improved. The desired height is between 1.60 and 1.70 m. The breeding goal weights the following traits as follows: conformation 40% and performance capability 60%. Performance includes soundness and temperament.
Oldenburg	A noble, generously lined, high performing sport horse with active impulsion and space gaining, elastic movements which, because of its predisposition, is permanently suitable for any type of sport (dressage, show jumping, eventing, driving).
Selle Français	Horses performing in competition with good conformation, gaits and temper. Priority is the jumping competition; but also eventing. Effort is made for dressage. There is also a highly performing branch of SF studbook specialised in steeple chase.
Sport Horse Breeding of Great Britain	An athletic, good moving, sound rideable sports horse of correct conformation that can compete in a range of disciplines, including eventing, dressage and show jumping and meets the needs of both the British rider and the wider international market.
Italian Saddle Horse	High quality sport horses particularly suitable for jumping.
Swedish Warmblood	A noble, correct and durable sport horse which through its temperament, rideability, good movements, and/or jumping ability is internationally competitive in dressage, show jumping or eventing.
Trakehner	Sound horse with Trakehner type, a big frame and correct and harmonious body proportions. Versatile riding and performance horses, easy to ride, with an energetic, elastic and ground-covering way of moving. Good and stable temperament, spirited but kind, intelligent, very willing to perform and with a tremendous stamina. All colours, usually standing between 1.60 and 1.70 m.
Westphalia	The Westphalian breeders should have the possibility to sell horses of every age-class (foal, mare, stallion, young riding horse, tested riding horse) for all disciplines. Important selection points include conformation, basic paces, riding ability and jumping potential.

Koenen, E.P.C., Aldridge, L., and Philipsson. (2004) "An overview of breeding objectives for Warmblood sport horses ", Livestock Production, 88, 77-84.

1b. Levels of government involvement in the horse industries

		Level of government involvement		
		Low	Moderate	High
Nature of horse industry	Fragmented	England USA[4] New Zealand		
	Semi structured		Ireland Canada Denmark	
	Organised	Netherlands		France Germany Sweden

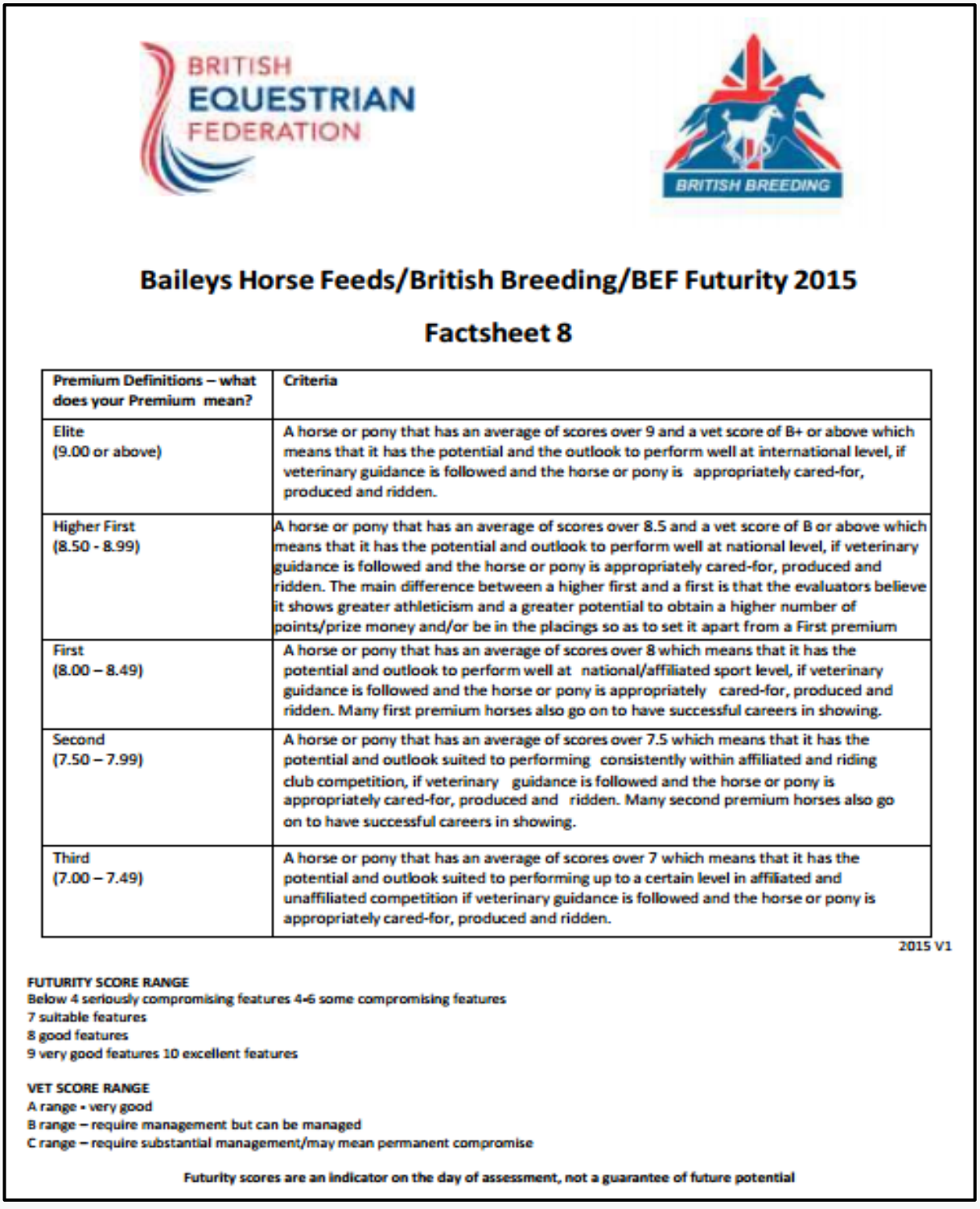
Crossman, G.K., (2010) "The organisational landscape of the English horse industry: a contrast with Sweden and the Netherlands", University of Exeter PhD, [online], https://ore.exeter.ac.uk/repository/bitstream/handle/10036/111475/CrossmanG.pdf, date accessed: 20.10.15.

1c. Futurity eligibility

Summary of Futurity eligibility	
British Bred	All entries must be British bred. British bred means that the horse was either foaled in Great Britain or foaled abroad to a dam normally living in Great Britain that had been temporarily exported for the purpose of breeding only. Neither the country of origin of the dam nor the country where the foal's sire is based nor the country of issue of the horse's passport is relevant in determining whether or not the foal is British bred. It is preferable that all entries are registered with a UK studbook. For further details see: http://www.bef.co.uk/repository/downloads/Horses/Futurity/2014/F5ELIGIBILITY2014V2.pdf
Age	In the Futurity, horses and ponies from foals to three year olds are eligible to enter

British Equestrian Federation Futurity: Factsheet 5, (2015b). [Online], http://www.bef.co.uk/repository/downloads/Horses/Futurity/2015/Futurity_Factsheet_5_FV.pdf, date accessed 20.10.15.

1d. BEF Futurity premium definitions 2015



BRITISH EQUESTRIAN FEDERATION

BRITISH BREEDING

Baileys Horse Feeds/British Breeding/BEF Futurity 2015

Factsheet 8

Premium Definitions – what does your Premium mean?	Criteria
Elite (9.00 or above)	A horse or pony that has an average of scores over 9 and a vet score of B+ or above which means that it has the potential and the outlook to perform well at international level, if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden.
Higher First (8.50 - 8.99)	A horse or pony that has an average of scores over 8.5 and a vet score of B or above which means that it has the potential and outlook to perform well at national level, if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden. The main difference between a higher first and a first is that the evaluators believe it shows greater athleticism and a greater potential to obtain a higher number of points/prize money and/or be in the placings so as to set it apart from a First premium
First (8.00 – 8.49)	A horse or pony that has an average of scores over 8 which means that it has the potential and outlook to perform well at national/affiliated sport level, if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden. Many first premium horses also go on to have successful careers in showing.
Second (7.50 – 7.99)	A horse or pony that has an average of scores over 7.5 which means that it has the potential and outlook suited to performing consistently within affiliated and riding club competition, if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden. Many second premium horses also go on to have successful careers in showing.
Third (7.00 – 7.49)	A horse or pony that has an average of scores over 7 which means that it has the potential and outlook suited to performing up to a certain level in affiliated and unaffiliated competition if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden.

2015 V1

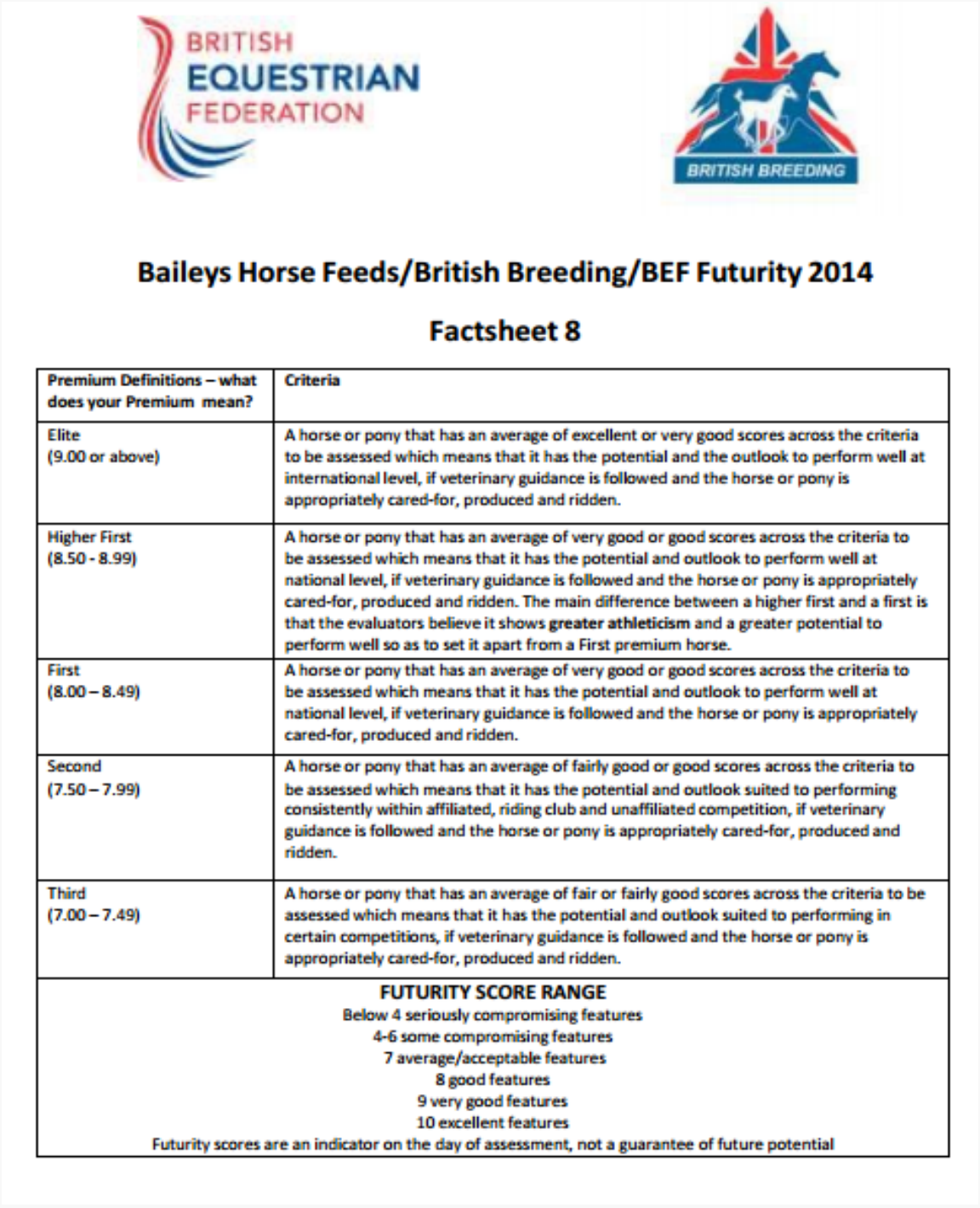
FUTURITY SCORE RANGE
Below 4 seriously compromising features 4-6 some compromising features
7 suitable features
8 good features
9 very good features 10 excellent features

VET SCORE RANGE
A range - very good
B range – require management but can be managed
C range – require substantial management/may mean permanent compromise

Futurity scores are an indicator on the day of assessment, not a guarantee of future potential

British Equestrian Federation Futurity: Factsheet 8, (2015c). [online], http://www.bef.co.uk/repository/Factsheet_8.pdf, date accessed: 22.10.15.

1e. BEF Futurity premium definitions 2014



Baileys Horse Feeds/British Breeding/BEF Futurity 2014

Factsheet 8

Premium Definitions – what does your Premium mean?	Criteria
Elite (9.00 or above)	A horse or pony that has an average of excellent or very good scores across the criteria to be assessed which means that it has the potential and the outlook to perform well at international level, if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden.
Higher First (8.50 - 8.99)	A horse or pony that has an average of very good or good scores across the criteria to be assessed which means that it has the potential and outlook to perform well at national level, if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden. The main difference between a higher first and a first is that the evaluators believe it shows **greater athleticism** and a greater potential to perform well so as to set it apart from a First premium horse.
First (8.00 – 8.49)	A horse or pony that has an average of very good or good scores across the criteria to be assessed which means that it has the potential and outlook to perform well at national level, if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden.
Second (7.50 – 7.99)	A horse or pony that has an average of fairly good or good scores across the criteria to be assessed which means that it has the potential and outlook suited to performing consistently within affiliated, riding club and unaffiliated competition, if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden.
Third (7.00 – 7.49)	A horse or pony that has an average of fair or fairly good scores across the criteria to be assessed which means that it has the potential and outlook suited to performing in certain competitions, if veterinary guidance is followed and the horse or pony is appropriately cared-for, produced and ridden.
FUTURITY SCORE RANGE Below 4 seriously compromising features 4-6 some compromising features 7 average/acceptable features 8 good features 9 very good features 10 excellent features Futurity scores are an indicator on the day of assessment, not a guarantee of future potential	

British Equestrian Federation: BEF Futurity, Factsheet 8, (2014a), [online], http://www.bef.co.uk/repository/downloads/Horses/Futurity/2014/Factsheet_8_2014.pdf, date accessed 09.10.14.

1f. BEF Futurity scoring and feedback sheet 2014

BRITISH EQUESTRIAN FEDERATION

Futurity EVENTING Score and Feedback Sheet 2014

Name_____ Number_____

Criteria					
Veterinary mark (V) 10 (increments on scale of .25) For limbs, hooves and musculoskeletal frame. The vet's mark will be an average (Av) of the score they attribute today and the score they believe the horse would obtain when competing at maturity provided any veterinary advice is followed (1 & 2). Vet score explanation: Below 4 Some serious compromising features, 4 - 6 Some compromising features which will require management 7 Mainly good features, minor management may be required 8 Good features 9 Very good features, 10 Excellent features	Comments:		1	2	Av
Frame (F) and build for EVENTING 10 Should have a rectangular build with horizontal back and proportional leg length. Supple poll and head/neck connection with clean throat latch. Good neck length, arched, neither set too high nor too low, with muscling to top line rather than underneath. Appropriately muscled back and loin with good wither, shoulder and saddle position. Should give a proportional, balanced impression, standing over sufficient ground with well positioned limbs.	Comments:	1	2	3	Av
Correctness of paces Walk for EVENTING (CPW) 10 Should show a 4-time gait which is active, straight, rhythmical and has impulsion and purpose. Should show suppleness through the back and body, a swinging stride, freedom in the shoulder, elbow, hip and stifle and a noticeable over track. Allowances are made for a foal under 21 weeks	Comments:				
Correctness of paces Trot for EVENTING (CPT) 10 Should show a 2-time gait which is active, straight, rhythmical and has impulsion and purpose. Should be supple through the body with balance and self carriage on the straight and on the turns. A "rounder" action is preferred to a straight one and there should be a noticeable lightness of footfall.	Comments:				
Correctness of paces Canter and Gallop for Eventing (CPC) 10 Canter: Horse shows a 3-time gait which is active, straight, rhythmical and has impulsion and purpose. Shows suppleness through the back and body and balance and self carriage on the straight and on the turns/direction changes/transitions. Should be able to lengthen and shorten its stride without loss of rhythm or balance. There should be a lightness of footfall and the body should remain horizontal or have a slight rise/lift in front. A good jumper will have a good canter. Gallop: This should be free, light, forward-going and ground covering. Should be a noticeable difference in pace between the canter and the gallop and balanced upward and downward transitions, the body should be lowered when travelling at speed but footfall should not be excessively heavy.	Comments:				
Athleticism (A) for EVENTING 20 Should show good joint flexion and effective use of the body to balance and carry the horse through turns and direction changes. Desire to move forward in all paces without excessive encouragement. Suppleness through the topline, without inversion and excessive tension.	Comments:				
Jump (J) for three year olds only (10) Able to collect in the final canter stride before the jump but can also take off on a long stride when required. Places hind legs well underneath the body in preparation for take off which should have thrust. Jumps with an uplift in the wither but may show a flatter bascule than a showjumper. Quickness of reflex to draw the forearm horizontally and fold the cannon bone back under the forearm. Tucks the hindlegs and draws them up and away from the fence in the latter part of the bascule. Lands lightly and canters away easily and freely. Is careful, efficient and has real scope. If the horse makes a mistake when jumping, tackles a fence differently next time showing an ability to quickly assimilate information and self corrects if the stride is wrong. Goes forward boldly and willingly down the jumping lane, if tightness in the back or tension exist, it dissipates, horse focuses on the fences and uses energy efficiently and with purpose.	Comments:				
PREMIUM = Av of V + F + CPW + CPT + CPC + A (+ J) divided by 7 (or 8) Scores out of ten using increments of .25 and the following range of scores: Below 4 Serious compromising features, 4 - 6 Some compromising features, 7+ Average/acceptable features, 8+ Good features, 9+ Very good features, 10 Excellent features					

British Equestrian Federation Futurity: Factsheet 4, (2014b), [online], http://www.bef.co.uk/repository/downloads/Horses/Futurity/2014/Factsheet_4_Eventing_2014.pdf, date accessed: 28.12.14.

Permission to reproduce all Futurity Factsheets has been granted by Jan Rogers, Head of equine development, BEF.

1g. BEF Futurity equine bridge

Summary of the Futurity Equine Bridge

Aim	The aim of the Equine Bridge is to provide direction and support to riders and trainers and to help owners and breeders realise the best of their horse's ability.
Selection	Horses and ponies in the dressage, eventing and Showjumping sections of Futurity, which were foaled in 2010 and which achieved a Futurity score of 8.5 or above as a 3 year old have been invited.
Assessment	All horses will need to walk, trot and canter on both reins and work on a 20m circle. Showjumpers and Eventers will need to willingly jump a short course of around 85cm + (smaller for ponies) with some natural fences and the opportunity to show gallop for eventers. Horses will be assessed in hand on a hard surface by a vet after the ridden assessment.

British Equestrian Federation Futurity: Futurity Equine Bridge (2014c), [online], http://www.bef.co.uk/repository/downloads/Horses/Bridge/Futurity_Equine_Bridge_2014__FV.pdf, date accessed: 20.10.15.

Section 2. Methodology

2a. BEF Futurity questionnaire

British Equestrian Federation Futurity Questionnaire

Introduction

Thank you for taking the time to complete my brief survey. As part of my research masters, I am investigating the effectiveness of the British Equestrian Federation's Futurity program in assessing young sports horses. I am carrying out this research by comparing retrospective young horse Futurity scores with adult horse competition results in Dressage, Showjumping and Eventing. From preliminary statistical analysis, I have found that a large majority of Futurity evaluated horses are not going on to compete in the affiliated disciplines that they were evaluated for. The aim of this questionnaire is to determine the reasons for this large discrepancy between young horse Futurity evaluations and later adult horse competition participation and performance.

Note: This questionnaire is solely for individuals who have had a horse evaluated by the Futurity between 2007 and 2013. This questionnaire is designed to record the details of one horse evaluation. If you have had more than one horse evaluated, I would appreciate if you could complete this questionnaire again separately, i.e. One questionnaire per horse.

Question 1; On how many occasions was your horse evaluated by the Futurity?

- 1
- 2
- 3
- 4

Question 2; What age was your horse at their **first** Futurity evaluation?

- Foal
- Yearling
- Two year old
- Three year old

Question 3; What year was your horse **first** evaluated in?

- 2007
- 2008
- 2009
- 2010
- 2011
- 2012
- 2013

Question 4; Which discipline was your horse **first** evaluated for?

- Dressage
- Eventing
- Show Jumping

Question 4a; What premium did you horse achieve at their **first** evaluation?

- Third 7-7.49
- Second 7.50-7.99
- First 8.00-8.49
- Higher First 8.50-8.99
- Elite 9+

Question 5; If you horse was evaluated more than once, what age was your horse at their **most recent** Futurity evaluation?

- Foal
- Yearling
- Two year old
- Three year old

(If you horse has only been evaluated once, please skip to question 8)

Question 6; What year was your horse **most recently** evaluated in?

- 2007
- 2008
- 2009
- 2010
- 2011
- 2012
- 2013

Question 7; Which discipline was your horse **most recently** evaluated for?

- Dressage
- Eventing
- Show Jumping

Question 7a; What premium did you horse achieve at their **most recent** evaluation?

- Third 7-7.49
- Second 7.50-7.99
- First 8.00-8.49
- Higher First 8.50-8.99
- Elite 9+

Question 8; Since Futurity evaluation(s), what discipline and competition levels has your horse competed in? (Select all that apply)

- None – Go to question 11
- Ridden showing
- Breeding
- Local Riding club
- Unaffiliated Dressage
- Unaffiliated Showjumping
- Unaffiliated Eventing
- **Other (please specify)**
- Affiliated Eventing
  - BE80T
  - BE90
  - BE100
  - Novice
  - Intermediate
  - Advanced+
- Affiliated Dressage
  - Preliminary
  - Novice
  - Elementary
  - Medium
  - Advanced Medium
  - Advanced+
- Affiliated Showjumping
  - 80cm
  - 90cm
  - 1m
  - 1.10m
  - 1.20m
  - 1.30m
  - 1.40m
  - 1.50m+

Question 9; thinking about your answer to question 8, was this the competition level you expected/wanted your horse to reach?

- Yes – Go to question 12
- No

Question 10; What do you believe prevented your horse advancing to higher levels?

(Select all that apply)

- No interest from the rider/owner to compete to higher competition levels
- Financial reasons
- Personal commitments
- Lack of suitable facilities
- Horse injury/health issue
- Horse sold at home
- Horse sold abroad
- Other (please specify)…….
- Lack of a suitable trainer
- Lack of a suitable rider
- Lack of a comprehensive plan
- Unsuitable horse temperament/behaviour
- Unsuitable horse conformation/locomotion
- Breeding

Go to question 12

Question 11; If your answer to question 8 was “none”, what was the reason for this?

(Select all that apply)

- No interest from the rider/owner to compete
- Financial reasons
- Personal commitments
- Lack of suitable facilities
- Lack of a suitable rider
- Lack of a suitable trainer
- Unsuitable horse temperament/behaviour
- Unsuitable horse conformation/locomotion
- Breeding
- Horse injury/health issue
- Horse sold at home
- Horse sold abroad
- Other (please specify)……

Question 12; In your personal opinion, do you believe your horse’s Futurity premium is an accurate representation of your horse’s performance ability?

- Yes
- No
- Don’t know

Question 13; What were your reasons/motivations behind having your horse evaluated by the Futurity?

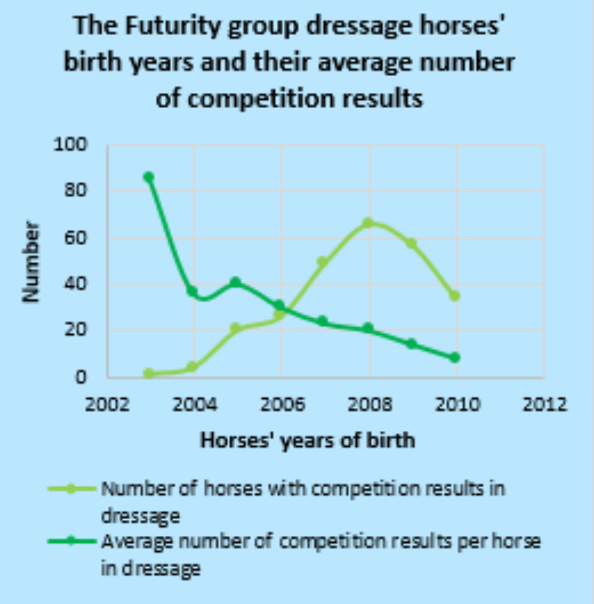
- To gain feedback on their potential suitability as a competition horse
- To gain feedback on their potential suitability as an amateur riding horse
- To potentially increase their monetary value
- To inform training decisions
- Other (please specify)…….

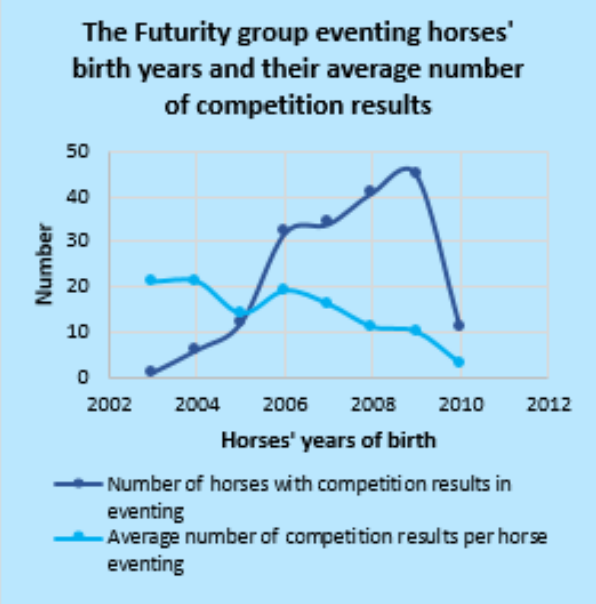
Section 3. Results

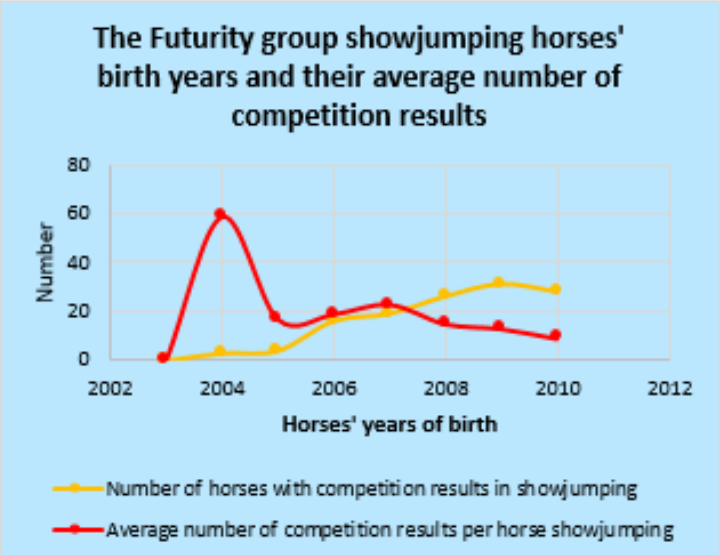
3a. Evaluation premiums awarded to Futurity group horses

Evaluation premiums awarded to Futurity group horses					
	Third 7-7.49	**Second 7.50-7.99**	**First 8.00-8.49**	**Higher first 8.50-8.99**	**Elite 9+**
Dressage	16	56	89	89	16
Eventing	8	37	74	47	13
Showjumping	1	22	58	34	11
Total	25	115	221	170	40

3b. The Futurity group horses' birth years and their average number of competition results







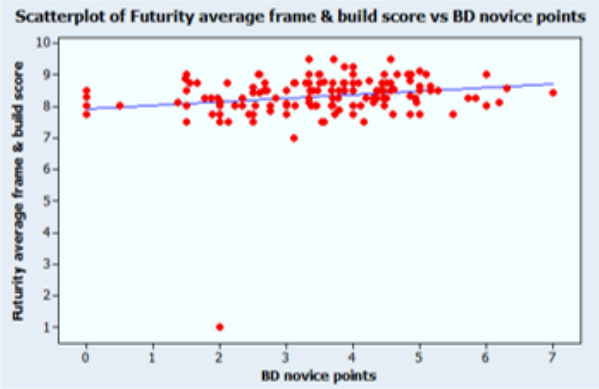
3c. Competition levels the futurity group horses competed in

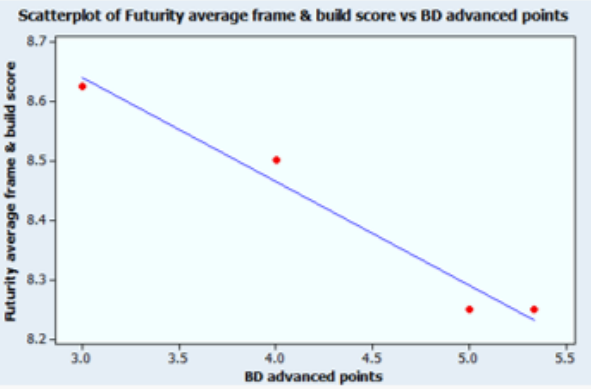
Competition levels the futurity group horses competed in								
	Preliminary	BDNovice BE80(T) BS90cm	BDEle'm BE90 BS1m	BDMed BE100 BS1.10m	BDAdMed BENovice BS1.20m	BDAdv BEInter BS1.30m	BDPSG BEAdv BS1.40m	BDInter BS1.50m
Dressage	257	229	117	61	21	8	4	2
Eventing		182	174	135	72	19	3	
Showjumping		127	92	61	29	11	4	2

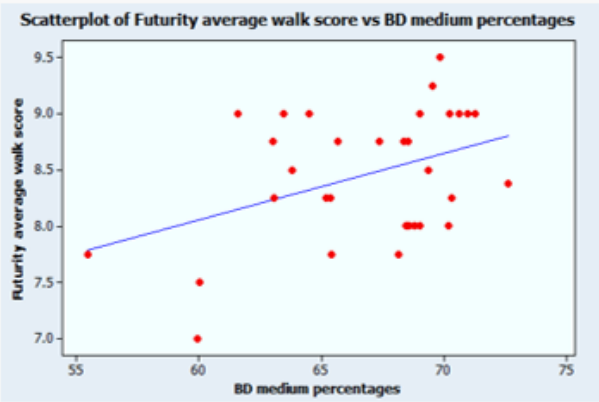
3d. Regression scatter plots (Futurity scores Vs performance measures)

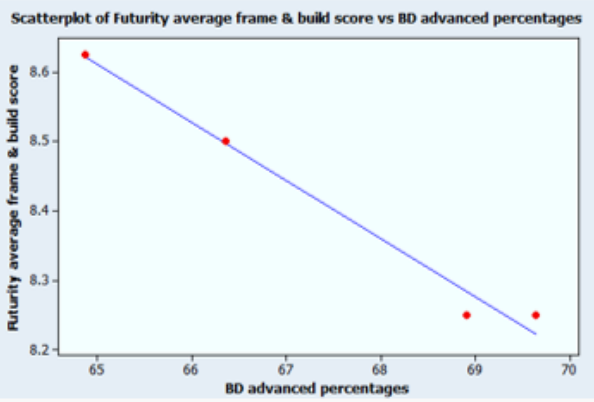
Table 15 scatter plots

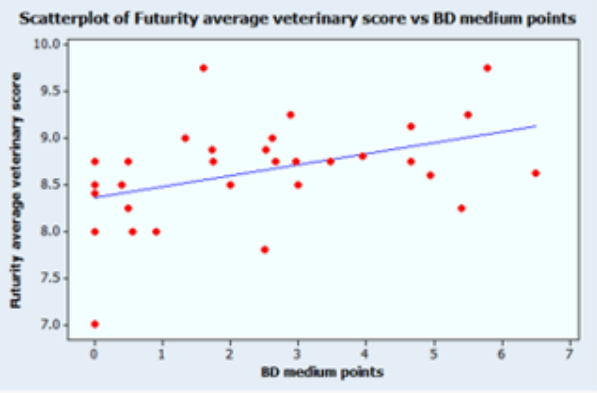
Performance measure: British dressage

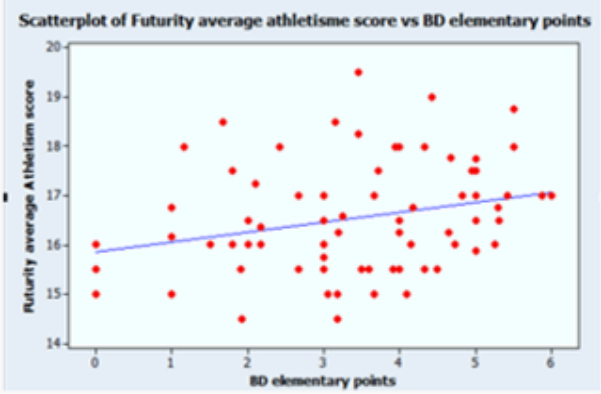


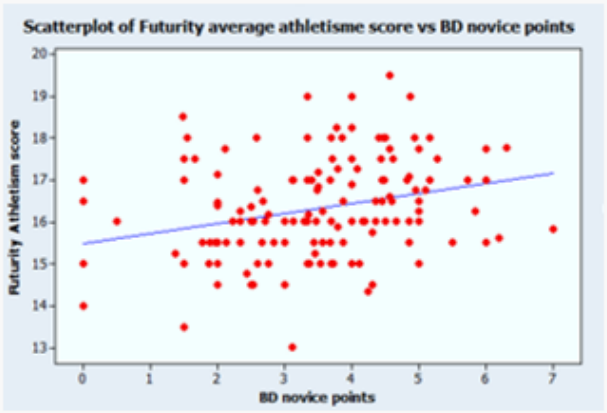


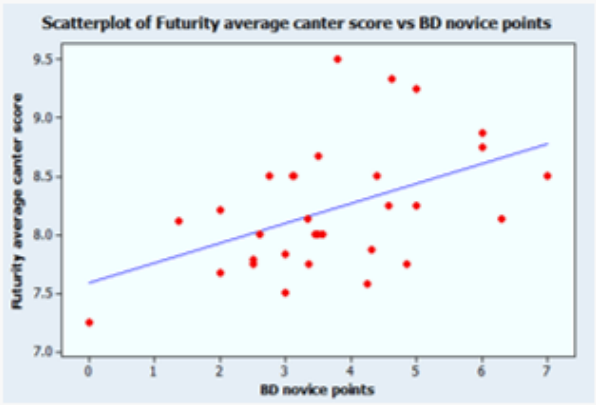


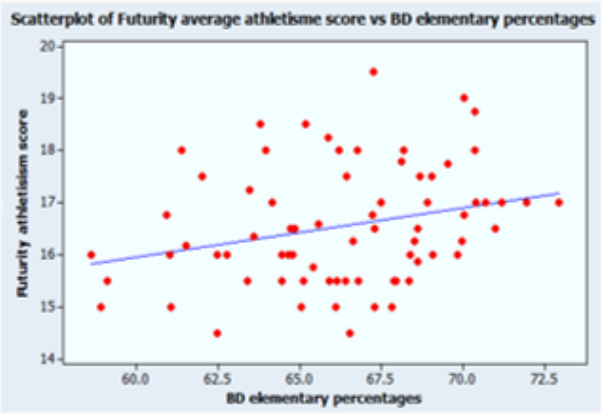


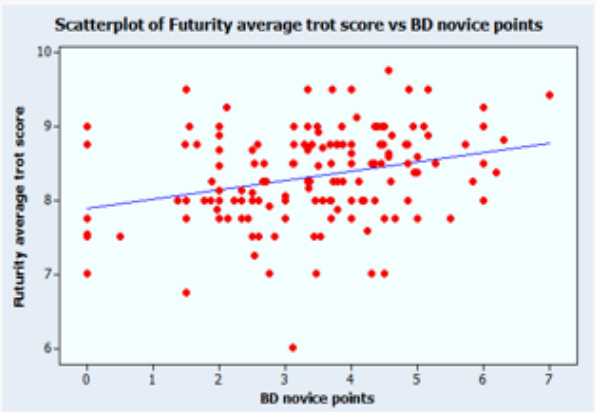


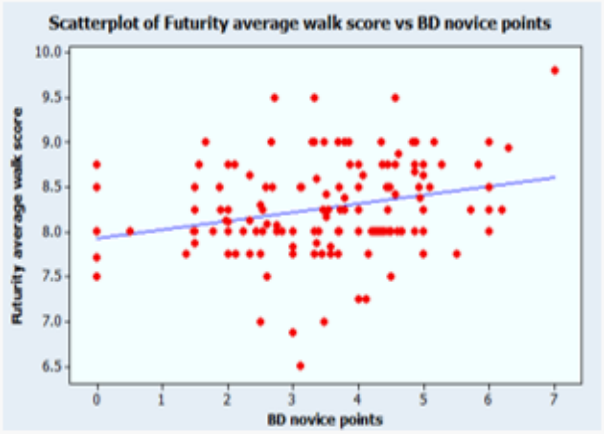


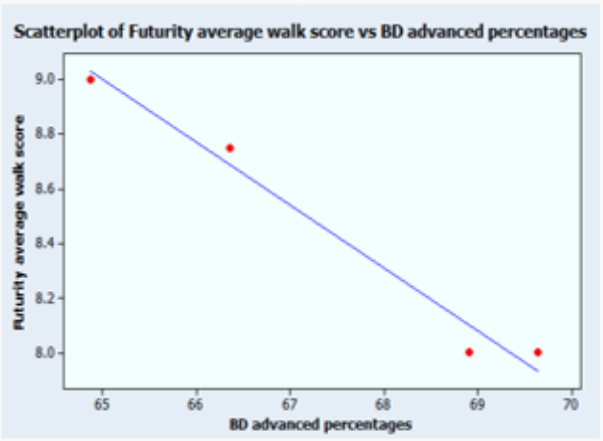

Scatterplot of Futurity average athletisme score vs BD novice points
Futurity Athletism score
BD novice points

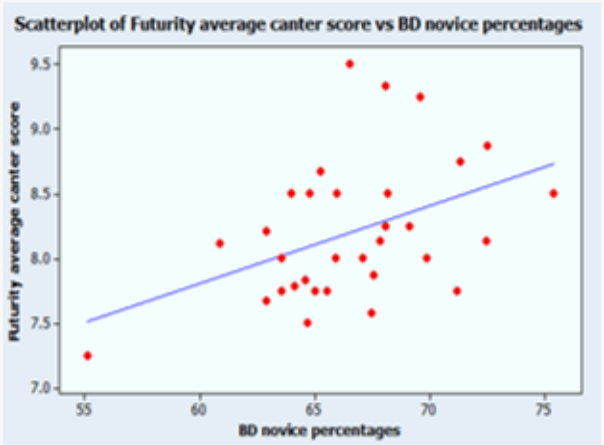

Scatterplot of Futurity average canter score vs BD novice points
Futurity average canter score
BD novice points

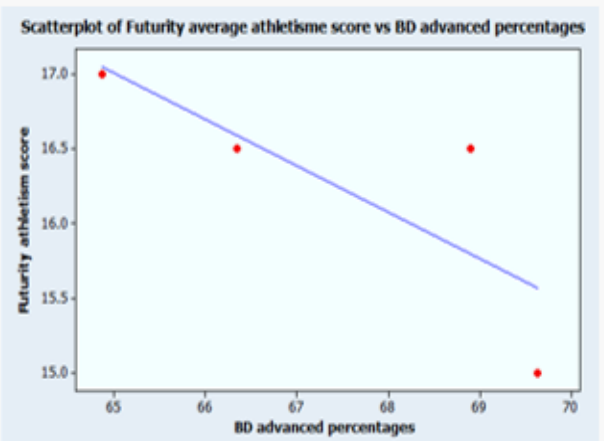

Scatterplot of Futurity average athletisme score vs BD elementary percentages
Futurity athletisism score
BD elementary percentages

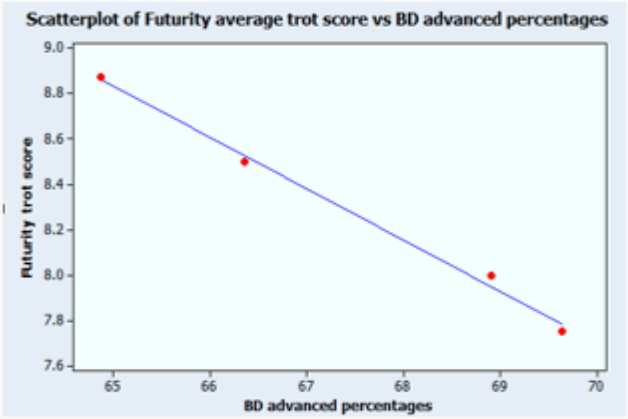

Scatterplot of Futurity average trot score vs BD novice points
Futurity average trot score
BD novice points

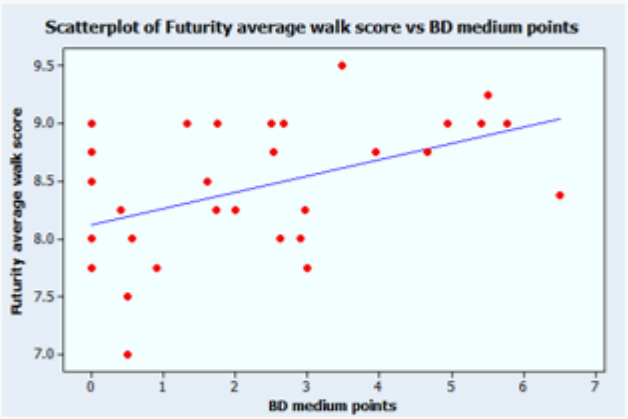

Scatterplot of Futurity average walk score vs BD novice points
Futurity average walk score
BD novice points

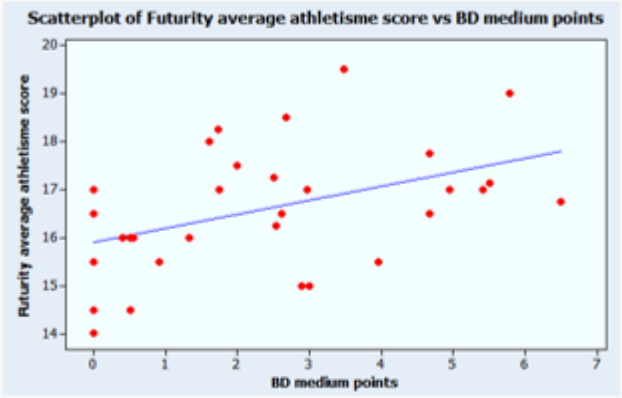

Scatterplot of Futurity average walk score vs BD advanced percentages
Futurity average walk score
BD advanced percentages


Scatterplot of Futurity average canter score vs BD novice percentages
Futurity average canter score
BD novice percentages

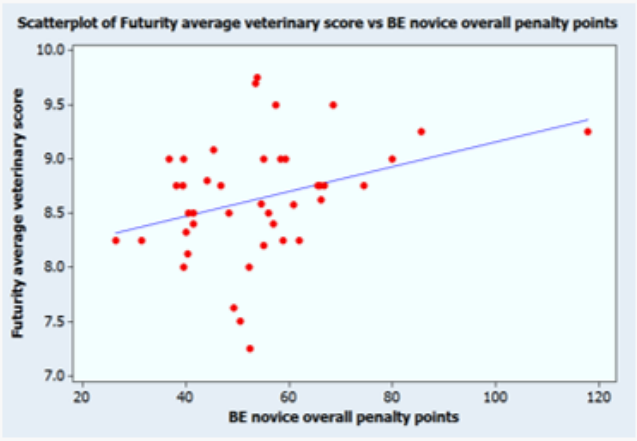

Scatterplot of Futurity average athletisme score vs BD advanced percentages
Futurity athletism score
BD advanced percentages

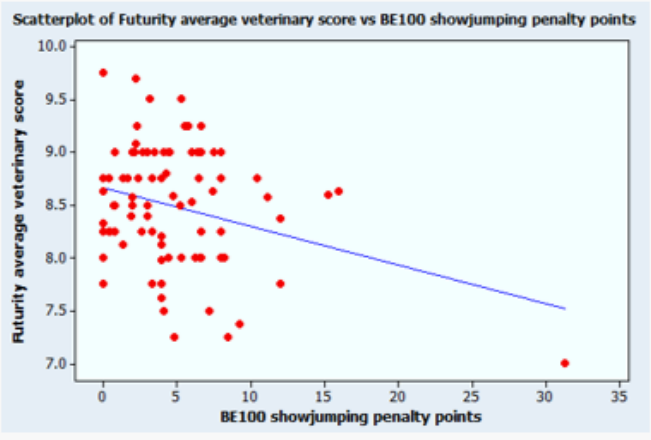


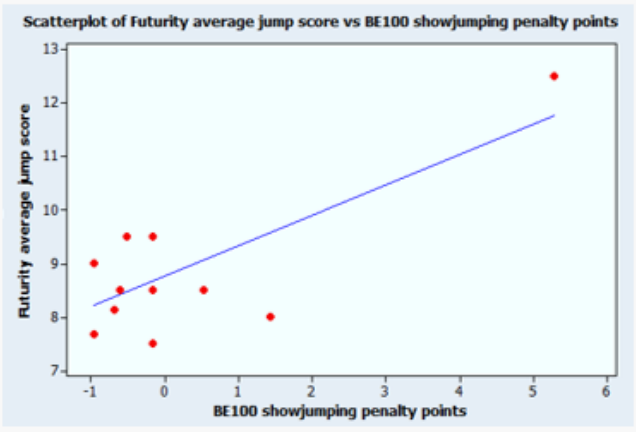


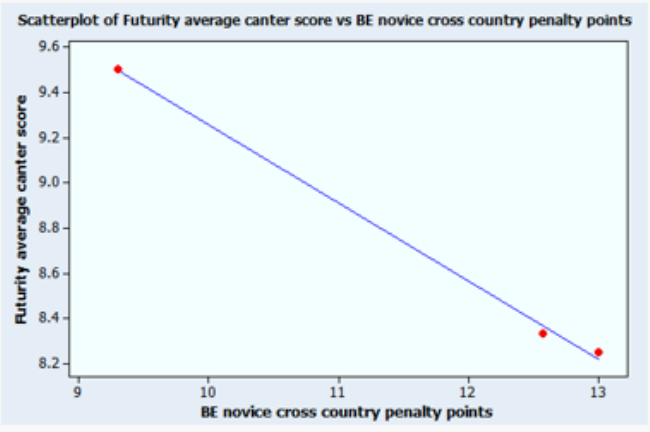


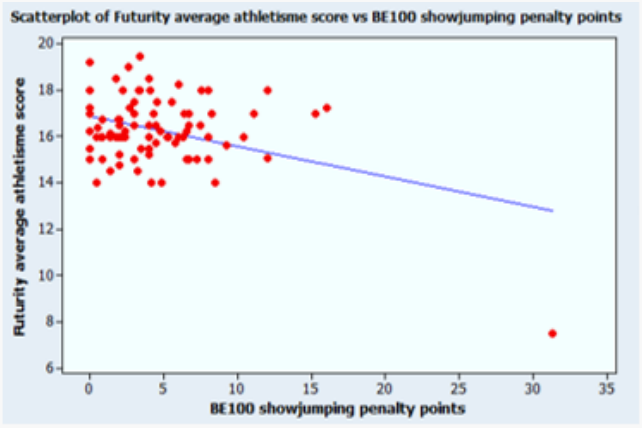
Performance measure: British eventing

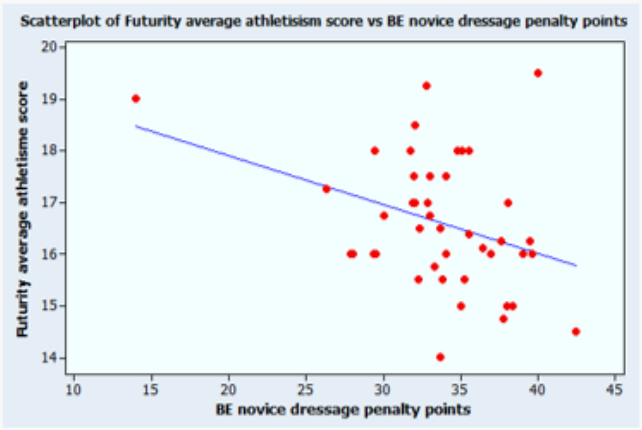


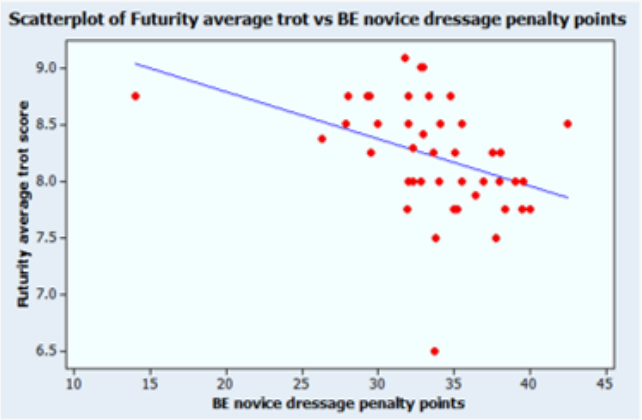


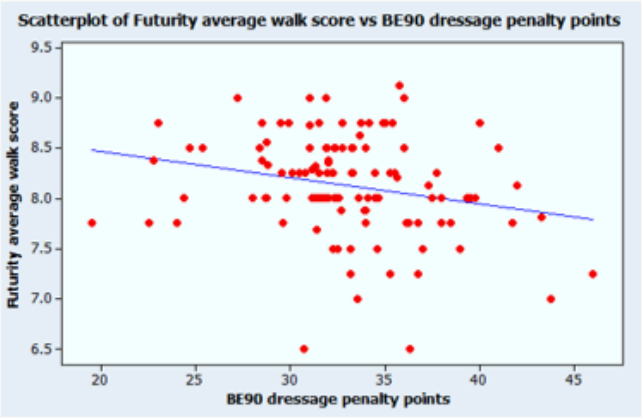


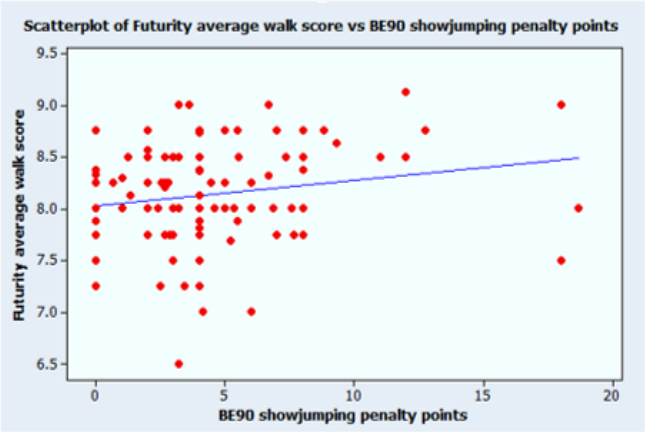




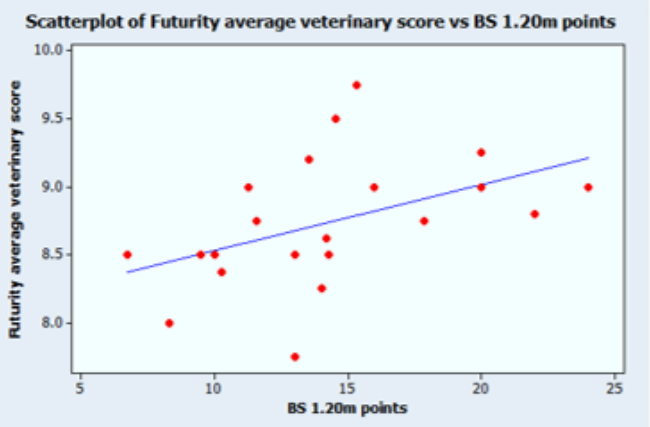


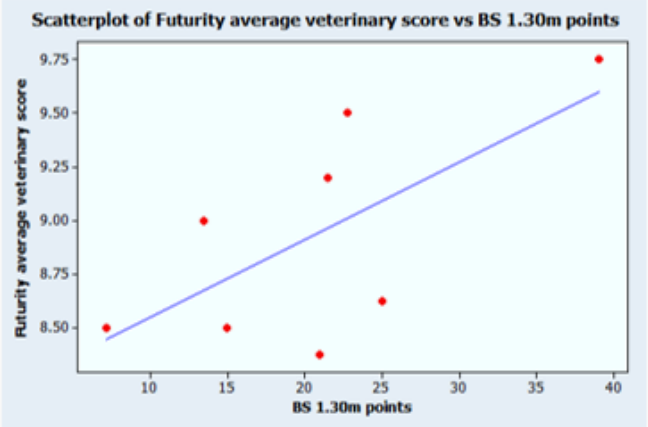


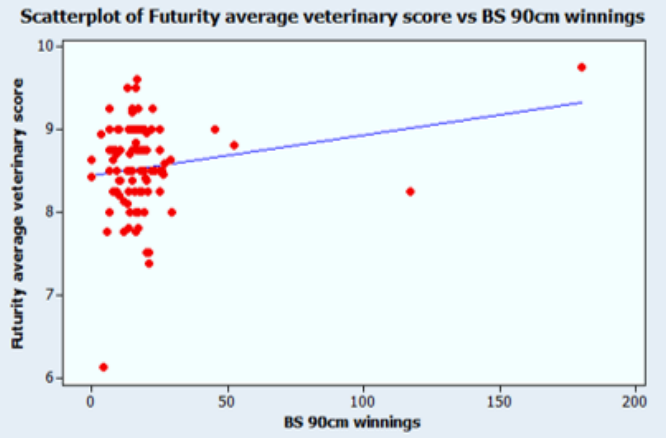


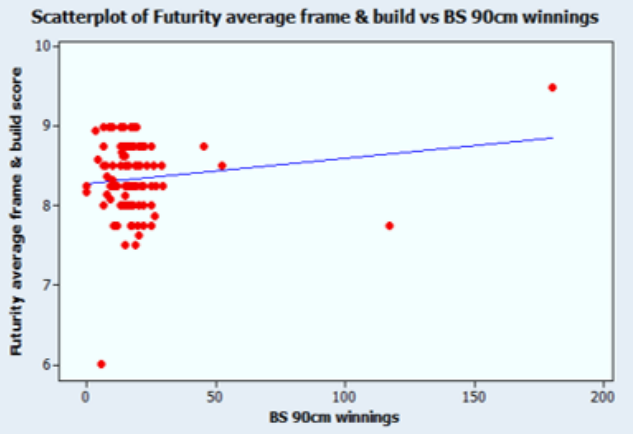


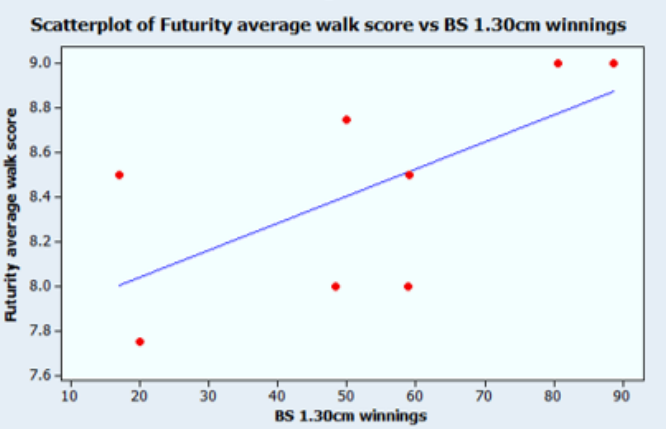
Performance measure: British showjumping

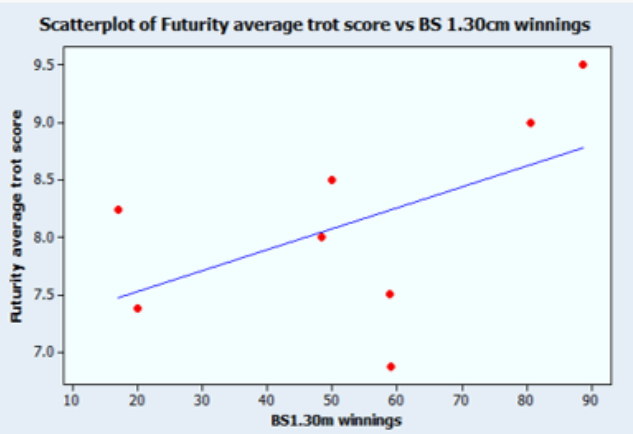


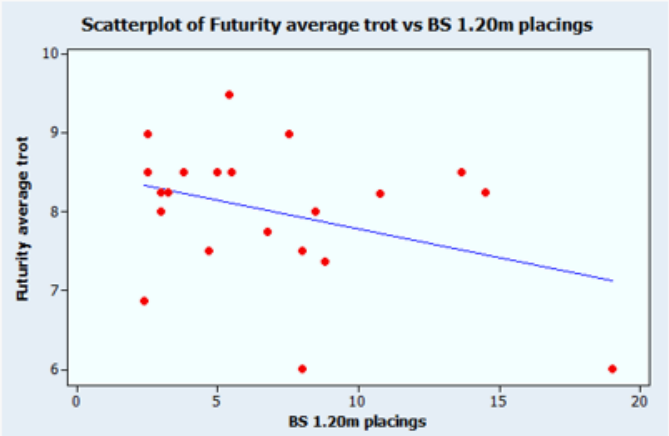



Scatterplot of Futurity average veterinary score vs BS 90cm winnings
Futurity average veterinary score
BS 90cm winnings


Scatterplot of Futurity average frame & build vs BS 90cm winnings
Futurity average frame & build score
BS 90cm winnings


Scatterplot of Futurity average walk score vs BS 1.30cm winnings
Futurity average walk score
BS 1.30cm winnings


Scatterplot of Futurity average trot score vs BS 1.30cm winnings
Futurity average trot score
BS1.30m winnings


Scatterplot of Futurity average trot vs BS 1.20m placings
Futurity average trot
BS 1.20m placings

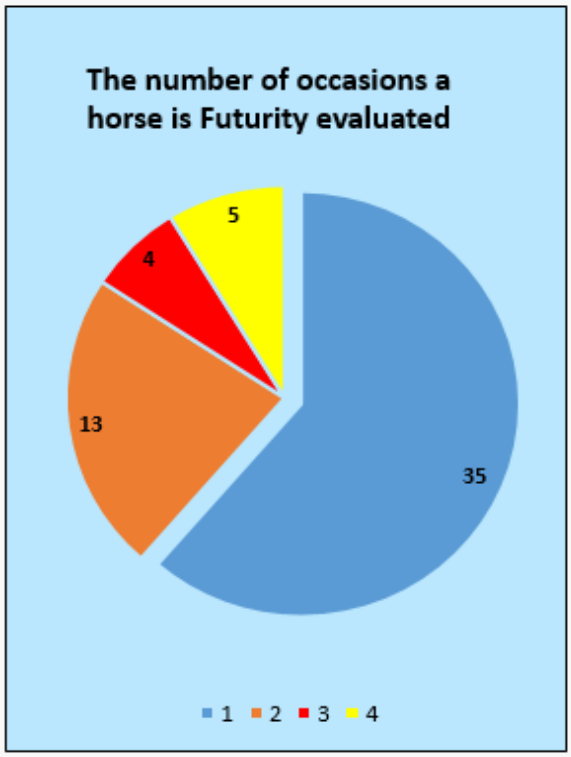
3e. Regression analysis of the Futurity group (R^2 values)

Regression analysis of the Futurity group		
Test: Averaged* performance measure (dependant variable) versus averaged* Futurity component scores (independent variable), Year of birth (categorical predictor)		
Discipline	**R^2**	**R^2 adjusted**
British Dressage	Component score: Veterinary mark	
BD medium - points	21.81%	10.23%
	Component score: Frame and build	
BD advanced - percent	99.95%	99.85%
BD novice - points	6.47%	3.25%
BD advanced - points	99.95%	99.85%
	Component score: Walk	
BD medium - percent	17.45%	5.22%
BD advanced - percent	99.65%	98.94%
BD novice - points	8.87%	5.73%
BD medium - points	29.24%	18.76%
	Component score: Trot	
BD advanced - percent	99.95%	99.85%
BD novice - points	9.44%	6.32%
	Component score: Canter	
BD novice - percent	21.84%	16.45%
BD novice - points	22.88%	17.37%
	Component score: Athleticism	
BD elementary - percent	8.7%	1.79%
BD advanced - percent	99.65%	98.94%
BD novice - points	10.28%	7.19%
BD elementary - points	7.28%	0.36%
BD medium - points	23.56%	12.24%
British Eventing	Component score: Veterinary mark	
BE novice – overall penalty points	22.29%	14.11%
BE100 – showjumping penalty points	10.56%	5.24%
BE novice – cross country penalty points	18.44%	9.85%
	Component score: Frame and build	
BE90 – dressage penalty points	6.04%	1.77%
	Component score: Walk	
BE90 - dressage penalty points	7.30%	3.09%
BE90 – showjumping penalty points	8.38%	4.21%
	Component score: Trot	
BE - novice dressage penalty points	18.12%	9.50%
	Component score: Canter	
BE novice – cross country penalty points	99.76%	99.52%
	Component score: Jump	
BE100 – showjumping penalty points	58.89%	49.75%
	Component score: Athleticism	
BE novice - dressage penalty points	17.62%	8.94%
BE100 – showjumping penalty points	16.49%	11.52%
British showjumping	Component score: Veterinary mark	
BS 1.20m – points	38.36%	17.81%
BS 1.30m - points	92.26%	81.94%
BS 90cm - winnings	11.29%	5.24%
	Component score: Frame and build	
BS 90cm - winnings	10.57%	4.47%
	Component score: Walk	
BS 1.30m - winnings	96.23%	91.21%
	Component score: Trot	
BS 1.20m - placings	47.40%	29.87%
BS 1.30m - winnings	96.64%	92.15%

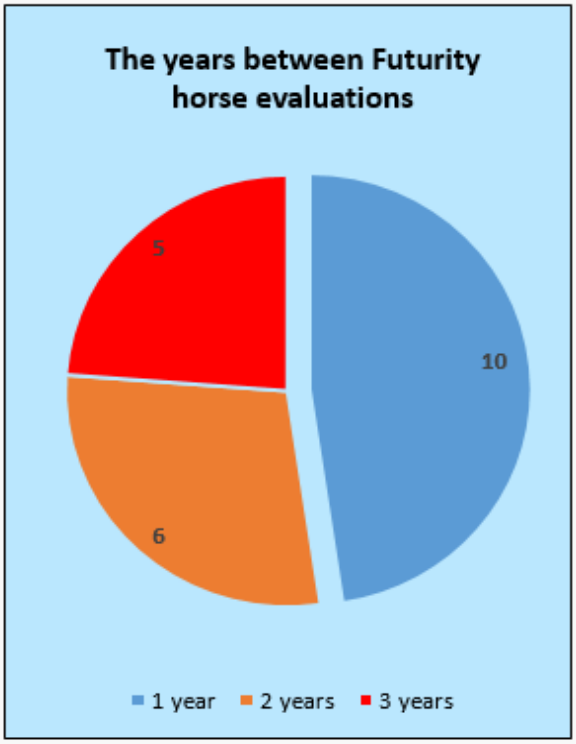
*Averaged scores were calculated by adding up the relevant values and then dividing by the number of numbers used
**Degrees of freedom

Section 4. Questionnaire results

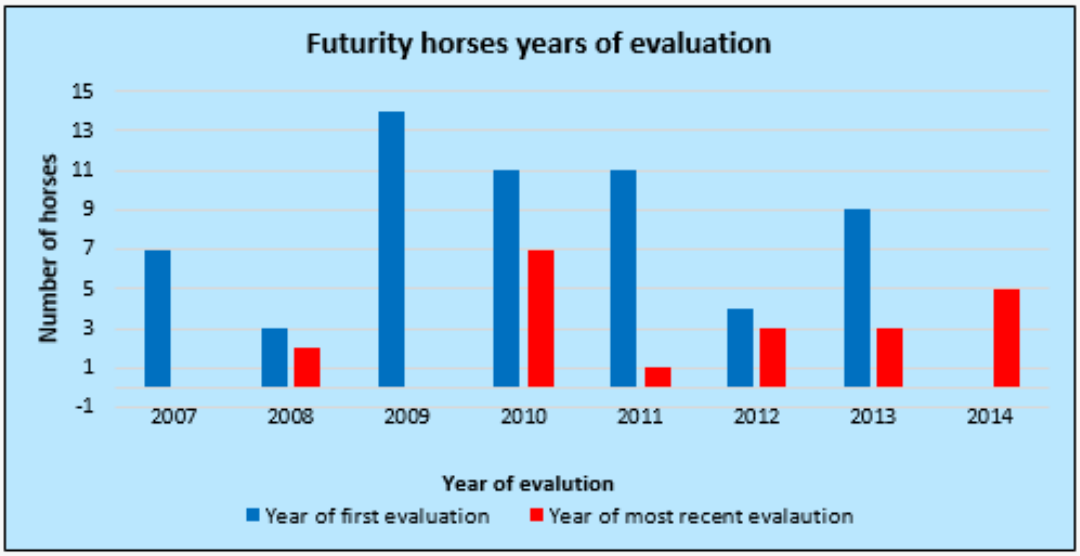
4a. **Number of occasions a horse is Futurity evaluated**



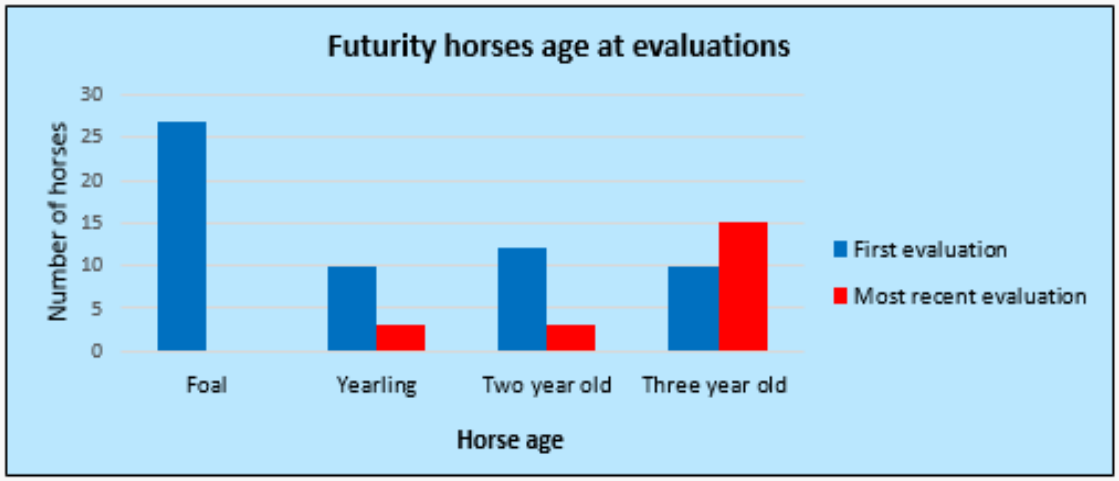
4b. **Years between Futurity horse evaluations**



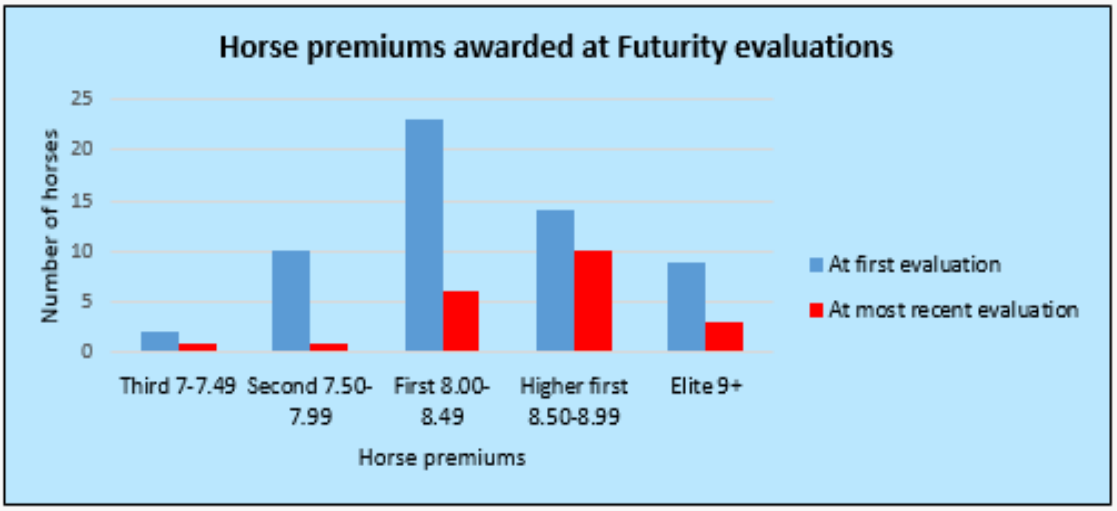
4c. Futurity horses years of evaluation



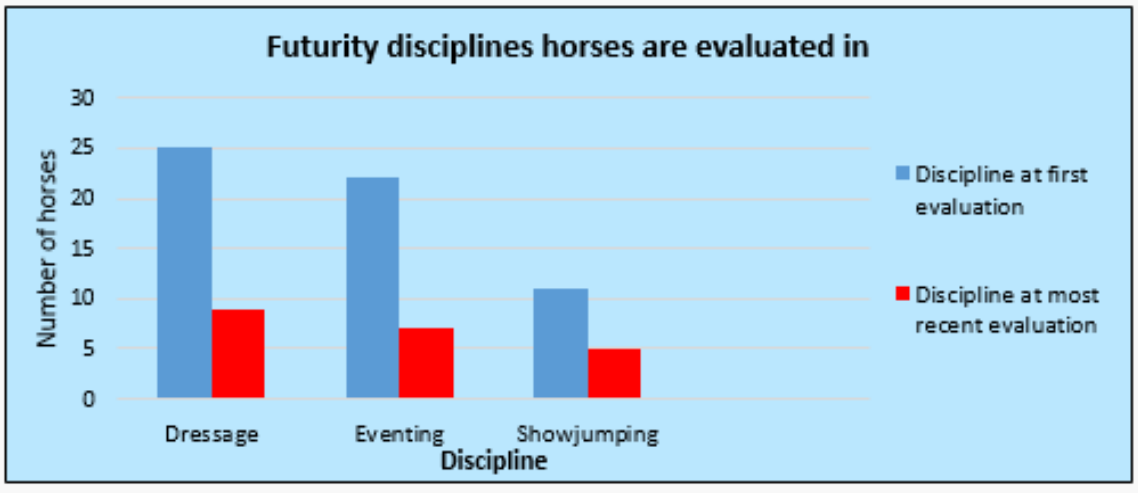
4d. Futurity horses age at evaluations



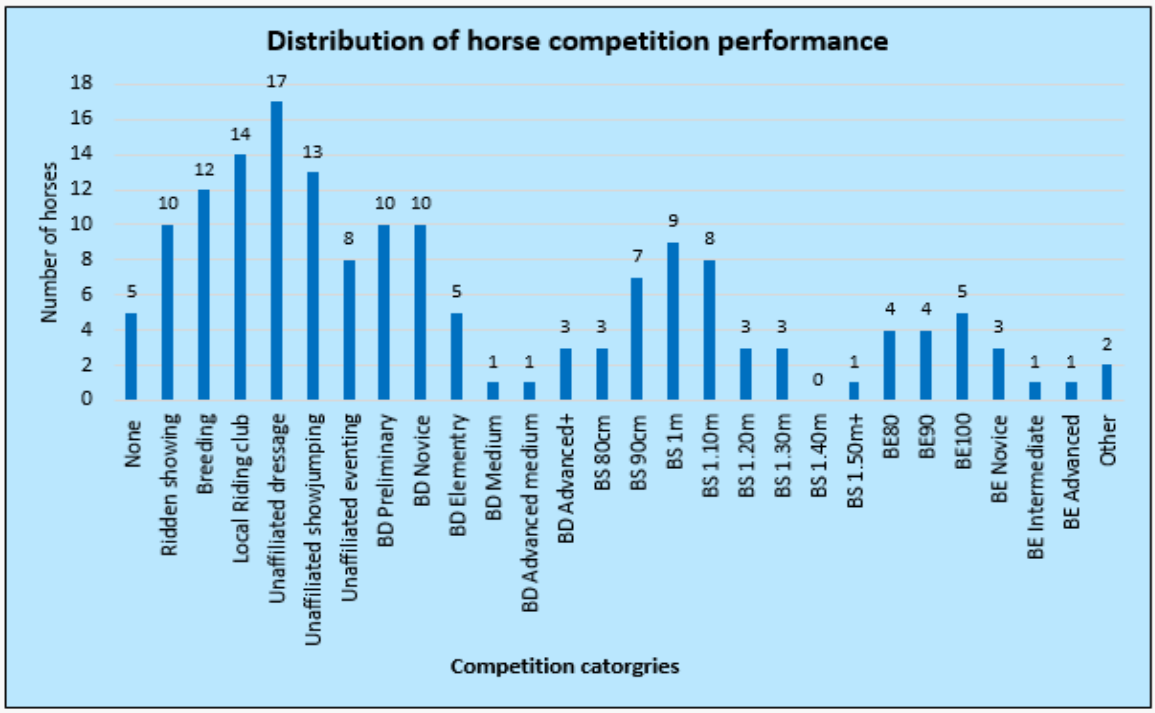
4e. Horse premiums awarded at Futurity evaluations



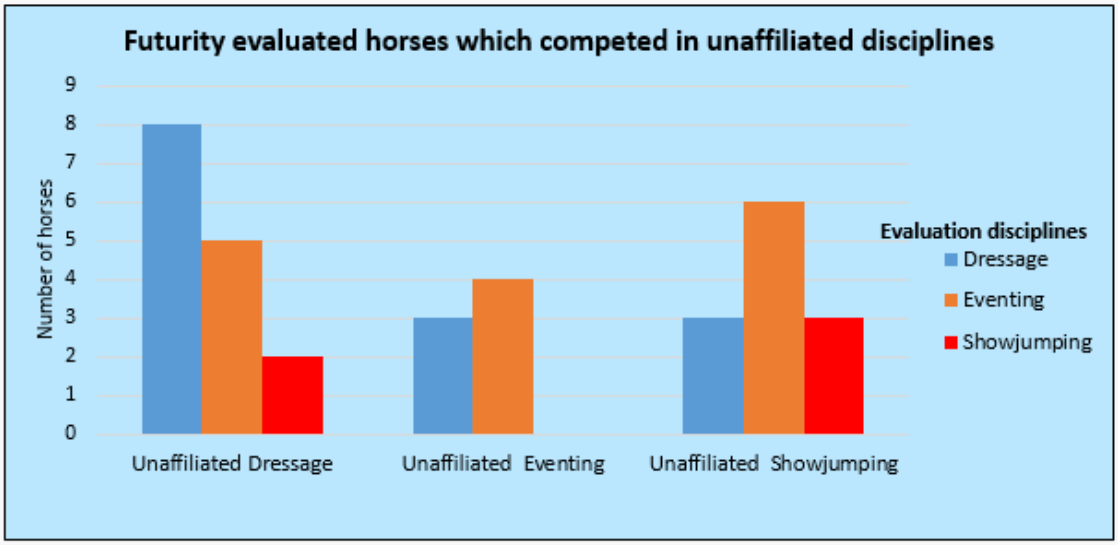
4f. Futurity disciplines horses are evaluated in



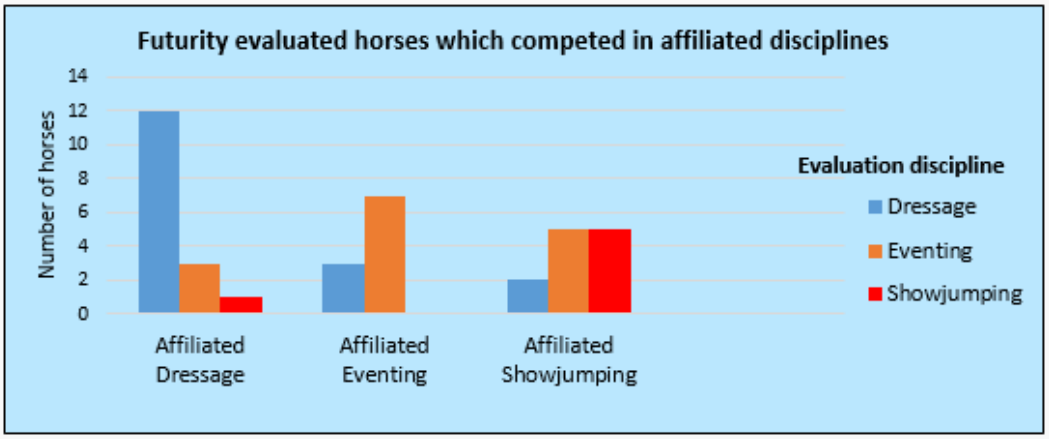
4g. Distribution of horse competition performance



4h. Futurity evaluated horses which competed in unaffiliated disciplines



4i. Futurity evaluated horses which competed in affiliated disciplines



4j. Futurity horse premiums and their affiliated competition performances

Futurity horse premiums and their affiliated BE competition performances						
Horse Premium	BE 80	BE90	BE100	BE Novice	BE Intermediate	BE Advanced
Third	-	-	-	-	-	-
Second	-	-	-	-	-	-
First	1	1	-	1	1	-
Higher First	1	-	1	-	-	-
Elite	-	-	2	-	-	1
Note: This table displays the most advanced level a horse reached						

Futurity horse premiums and their affiliated BD competition performances				
Horse Premium	BD Preliminary	BD Novice	BD Elementary	BD Advanced
Third	-	1	-	-
Second	2	-	-	1
First	2	3	2	2
Higher First	-	1	1	-
Elite	-	1	-	-
Note: This table displays the most advanced level a horse reached				

Futurity horse premiums and their affiliated BS competition performances				
Horse Premium	BS 1m	BS 1.10m	BS 1.30m	BS 1.50m
Third	-	-	-	-
Second	-	-	-	-
First	2	4	1	-
Higher First	1	2	1	-
Elite	1	-	1	1
Note: This table displays the most advanced level a horse reached				

4k. Chi-Square Test Results of horse premium scores and competition affiliation

Chi-Square Test 1 Results: Horse premium scores

	Did the horse affiliate?	
Premium groups	Yes	No
Third and Second	3	9
First and Higher Frist	24	13
Elite	7	2

Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

	Yes	No	Total
1	3	9	12
	7.03	4.97	
	2.314	3.278	
2	24	13	37
	21.69	15.31	
	0.246	0.349	
3	7	2	9
	5.28	3.72	
	0.563	0.798	
Total	34	24	58

Chi-Sq = 7.548, DF = 2, P-Value = 0.023
2 cells with expected counts less than 5.

Chi-Square Test 2 Results: Horse premium scores

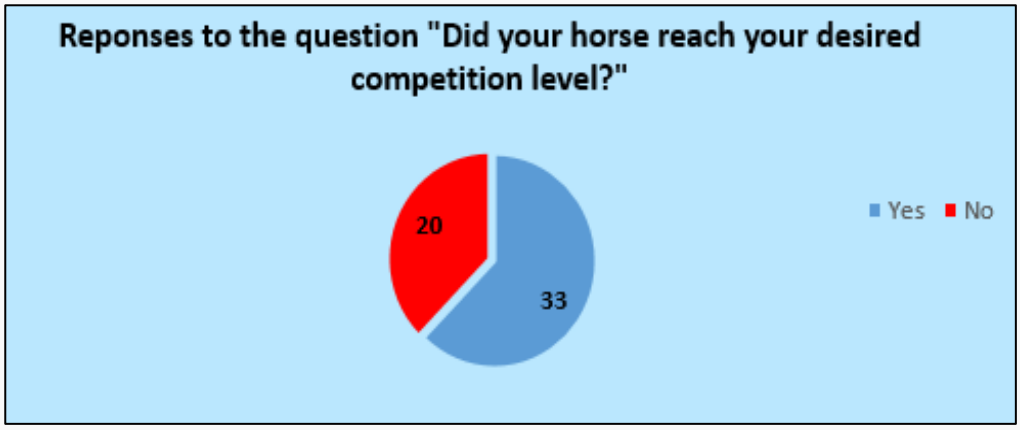
	Did the horse affiliate?	
Premium groups	Yes	No
Third and Second	3	9
First, Higher Frist and Elite	31	15

Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

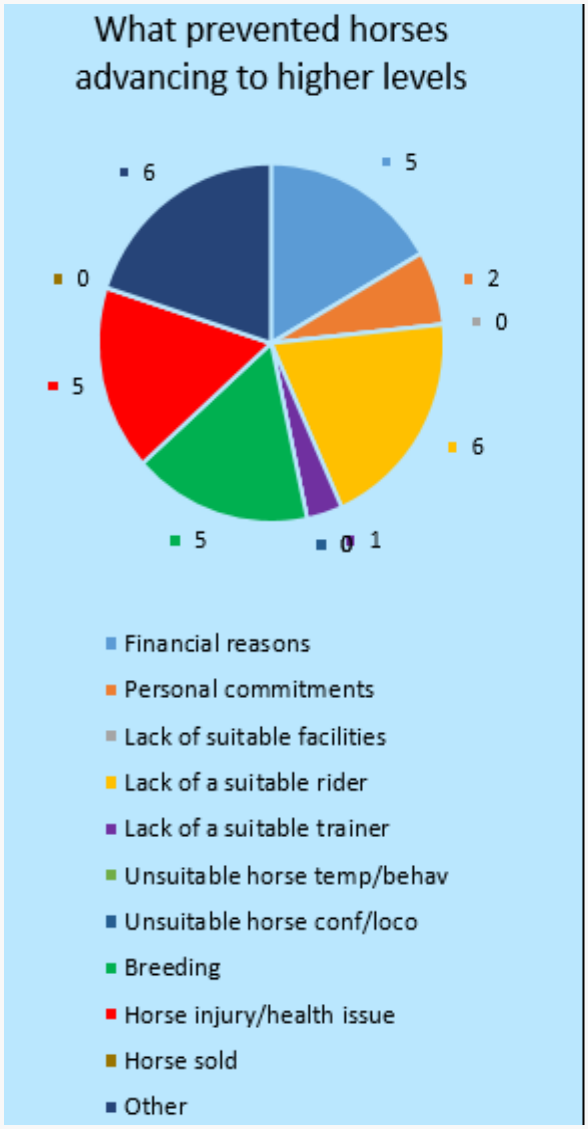
	Yes_2	No_2	Total
1	3	9	12
	7.03	4.97	
	2.314	3.278	
2	31	15	46
	26.97	19.03	
	0.604	0.855	
Total	34	24	58

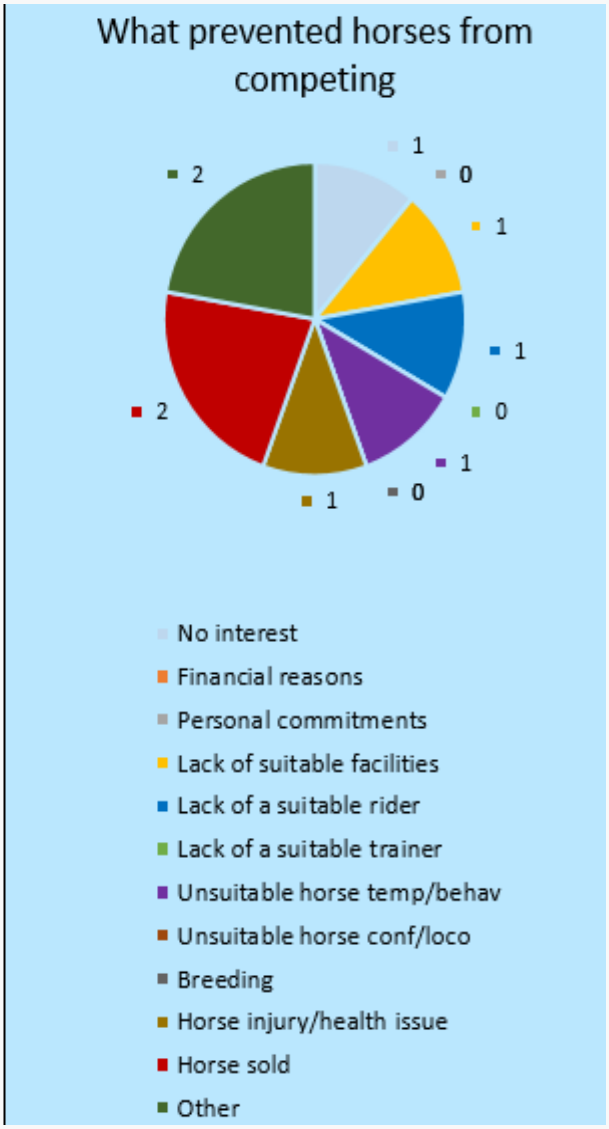
Chi-Sq = 7.051, DF = 1, P-Value = 0.008
1 cells with expected counts less than 5.

4l. Responses to the question “Did your horse reach your desired competition level?”

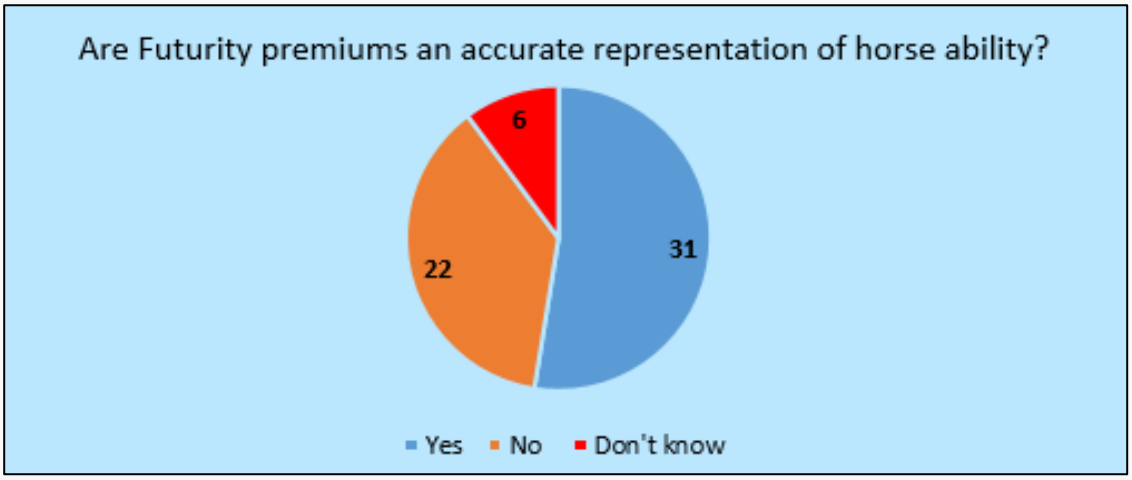


4m. What prevented horses competing and competing to advanced levels

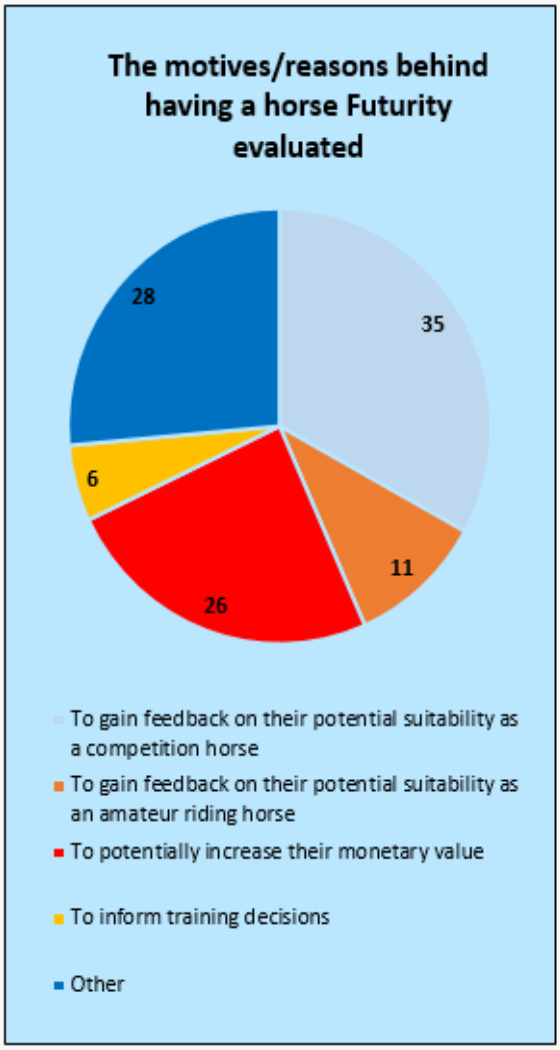


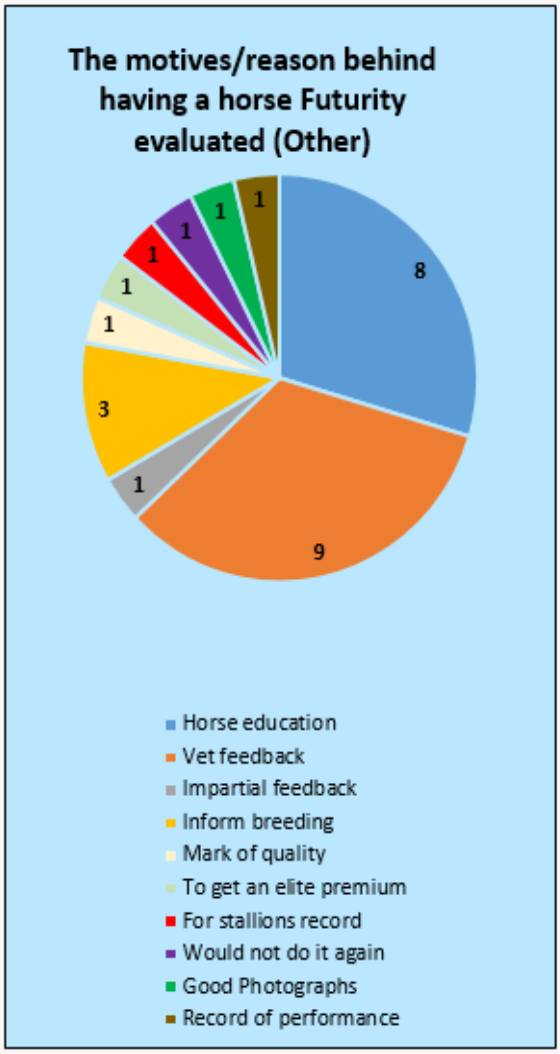


4n. Are Futurity premiums an accurate representation of horse ability?



4o. The reasons behind having a horse Futurity evaluated





Section 5. Researcher's credentials

5a. Credentials

Background research

Throughout the completion of the thesis, I attended various events to inform the progression of my research.

To experience the practical data collection of the origins of my data I volunteered at a BEF Futurity event at Bicton College in August 2014, and then with British Eventing in March 2015. Additionally in In November 2015 I attended the British Dressage offices to understand the data collection process and obtain my data. To understand the impact of equine research on the industry, I attended the British Society of Animal Science conference in March 2015 where Jan Rogers, of the BEF Futurity, was giving a talk on the progress on the programme.

In March 2016 I presented at the Breeding for Gold conference at the Royal Agricultural University in Cirencester. Jan Rogers invited me to give a talk alongside her at the conference regarding the preliminary results of my research.

Evidence for the above events are displayed below;

British Equestrian Federation Futurity Volunteer Evidence.

From: Cat [redacted]

Sent: 25 March 2015 10:28:15

To: Lauren [redacted]

Hi Lauren

Thank you for volunteering at the BEF/British Breeding Futurity event at Bicton College on Sunday 3rd

August 2014. Your help was very much appreciated and we hope you found it beneficial to your course.

Many Thanks

Cat

[redacted]

BEF Equine Development Coordinator
British Equestrian Federation

[redacted]

Aldon International Volunteer Evidence.

From: Tessa [redacted]

Sent: 27 March 2015 16:26:07

To: Lauren [redacted]

Hi Lauren,

This serves 2 purposes!

1. To provide evidence that you were at Aldon on 21st March,

2. To thank you so much for coming and assisting in the score wagon.

I do hope it gave you an insight into a small part of the event.

Many thanks again and good luck with your course.

Tessa [redacted]

Hon Organiser Aldon Horse Trials

Breeding for Gold conference evidence

From: Jan [redacted]

Sent: 10 March 2016 10:41:45

To: Lauren [redacted]

Dear Lauren,

Thank you very much for presenting your summary findings at the Breeding for Gold Symposium yesterday. The organisers were very grateful and very complementary about your style.

We are pleased to have been able to work with you on this project and would be very happy to continue to do so in ways as may become apparent.

Thank you very much once again.

Jan [redacted]

Head of Equine Development I British Equestrian Federation Equestrian House I [redacted]

BSAS Attendance Evidence

17 March 2015

BSAS Ref: 236/Student Non-member/Student

Miss Lauren

RECEIPT

Science with Impact
14-15 April 2015
University of Chester, CHESTER, , UK

Thank you for your application to attend the above meeting at the University of Chester. We acknowledge receipt of the following booking:

CONFERENCE FEE	**One Date Rate**	*£130.00*
ACCOMMODATION		
You are booked into accommodation on campus, B&B Ensuite		*£48.50*
Date of Arrival: ***14 April***	*Date of Departure:* ***15 April***	
DINNER		
Wednesday Tea/Coffee/Lunch -		
		£0.00
	TOTAL	**£178.50**

Paid in full with thanks. Attendance Wednesday.