Personality in young horses and ponies evaluated during breeding shows: phenotypic link with jumping competition results

Marianne Vidament, Léa Lansade, Sophie Danvy, Bernard Dumont Saint Priest, Margot Sabbagh, Anne Ricard

Personality in young horses and ponies evaluated during breeding shows: phenotypic link with jumping competition results

Marianne Vidament [a,b*], Lea Lansade [a,b], Sophie Danvy [b], Bernard Dumont Saint Priest [b], Margot Sabbagh [b] and Anne Ricard [b,c]

[a] INRAE, CNRS, IFCE, Université de Tours, MNHN, PRC, 37380 Nouzilly, France

[b] IFCE (Institut Français du Cheval et de l'Equitation), 49400 Saumur, France

[c] INRAE, UMR1313, Génétique Animale et Biologie Intégrative, 78352 Jouy en Josas, France

[*] Corresponding author: marianne.vidament@inrae.fr

Abstract

Animal personality, the result of temperament being modulated by life events, is an important factor to be considered when breeding and using domestic horses. In the breeding of sport horses, personality appears as a secondary trait in selection objectives after competition performance. Moreover, the personality trait of fearfulness may be viewed as a risk factor for riders. This study aimed to estimate the variability of personality characteristics measured during breeding shows and their phenotypic correlation with performance in jumping competitions. Data for personality characteristics were recorded during 67 breeding shows in France on 876 jumping horses, 424 jumping ponies and 45 leisure ponies aged 2 or 3 years. Their behavior was assessed during 1) customary rounds (CR) of breeding shows (conformation, free jumping and height measurement at withers) and 2) specific tests (ST) conducted in-hand that measured fearfulness (novel object, novel surface and suddenness tests) and tactile sensitivity. Not all the animals were evaluated on all the behavior tests. Jumping performances from 4 to 7 years old were recorded for 724 of the horses and for 313 of the ponies in official competitions specific for horses or ponies. Environmental effects were estimated using general linear model taking into account breeding show, age and sex. The breeding show effect was significant on 23 out of 28 characteristics. Age and sex influenced approximately one third of the characteristics: younger animals were more fearful; males moved and whinnied more; geldings appeared slightly more fearful during ST. Jumping performances were mostly independent of personality characteristics. In horses, performances were phenotypically positively linked with 3 characteristics during jumping CR (whinnies (P=0.05), main gait when entering (P=0.02), evasive behaviors (P=0.03)) and with posture during conformation evaluation CR (P= 0.04). In ponies, jumping performances were phenotypically positively linked only with whinnies: during CR of height measurement (P=0.02) and during all ST (P=0.01). As no main fear variables were significantly related to

jumping performances in the two studied populations, it seems that less fearful horses and ponies may perform well in show jumping.

Key words: personality, fearfulness, tactile sensitivity, horse, pony, show jumping performance

Introduction

Personality and temperament are defined as a set of individual behavioral attitudes which remain relatively stable across time and situations. Temperament is considered as the innate basis of these behavioral attitudes and personality is the result of the modulation of this temperament by life events (Feaver et al., 1986; Bates, 1989). Personality represents a major factor influencing the life of domestic animals and some temperament characteristics have a genetic basis, for instance, fearfulness in cattle or docility to milking in dairy cows (Burrow and Corbet, 2000; Canario et al., 2013; Haskell et al., 2014).

In horses, different personality traits have been studied: fearfulness, gregariousness, reactivity, learning ability, reactivity to humans, sensory sensitivities (see for review (von Borstel, 2013)). In fearful or stressful situations, horses can exhibit strong reactions compromising their welfare (Hall et al., 2018) and leading to human injuries (Wolyncewicz et al., 2018). Personality tests used in this study were developed in our laboratory (Lansade et al., 2008a, b, c). They were simplified to be usable in field conditions. They evaluate mainly fearfulness and tactile sensitivity (Lansade et al., 2016).

Less fearful horses make riding safer for riders (Visser E. K. et al., 2008; Rothmann et al., 2014; Lansade et al., 2016). However, riders may wonder if these less fearful horses will be as good in competitions as more fearful animals. On the other hand, high tactile sensitivity may be viewed as a quality for good rideability. The literature on the relationship between

personality and jumping abilities is limited and contradictory (Visser EK et al., 2003; Rothmann et al., 2014; Lansade et al., 2016). Therefore, knowledge of the phenotypic relationship between these personality traits and competitive performance is of great interest to riders.

This study aimed to evaluate factors influencing personality traits and the phenotypic correlations between these characteristics and jumping performance. The personality of young horses and ponies was measured using simplified personality tests conducted in the field during breeding shows.

Material and methods

The simplified personality tests consisted in observing and rating behavior during four specific tests (ST) conducted on horses and ponies held in hand and during the three customary rounds (CR) of breeding shows: height measurement, conformation assessment and free jumping. In this study, the term “animal(s)” will be used to refer to horses and ponies, except when the two are differentiated.

Animals and locations of breeding shows

A campaign to characterize the personality of horses and ponies on a large scale was run from 2013 to 2015 during 67 breeding shows in France. These breeding shows were open to ponies and horses of different breeds and had very similar rules. The judges were often the same in a given region. However, the height of obstacles was adapted to the size of the horses or ponies. Free jumping was not performed by leisure ponies.

The personality assessment (CR and ST) was similar for all breeds. If by chance an animal was tested in two breeding shows, only the results of the first show were kept. This study presented data of simplified personality assessments performed on a total of 1,345 animals: 682 in 2013, 372 in 2014 and 291 in 2015.

Not all animals were assessed for personality in all the behavior tests because of the following reasons. Firstly, according to the number of animals in the breeding shows, two to four assessors were needed because ST and CR took place simultaneously. If an assessor was missing for one particular CR or for the ST, animals were not assessed on that aspect in that breeding show. Secondly, all animals that were present were assessed in the different CR when organized, but only animals presented voluntarily by their owner or trainer were assessed on the ST. This proportion represented around 50-70% of animals presented at the breeding shows.

Therefore only 734-1,100 observations of individual characteristics were completed depending on the variables recorded (see details in Tables 2a and 2b). Animals were 2 and 3 years old (12% and 88% respectively). There were 715 females (53%), 347 geldings (26%) and 283 males (21%). They were grouped according to their breed into four types: 1) sport horses (SH, n = 810: mostly French Saddle horses n = 737, foreign sport horses n = 15 and horses originating from sport horses but not registered in a studbook n = 58, all these breeds are cross-breeds), 2) Anglo Arabians (AA, n = 66), 3) sport ponies (SP, n=424: French Saddle ponies, Connemara, Dartmoor, New Forest, Welsh C or D or originating from these breeds) and 4) leisure ponies (LP, n = 45: Haflinger, Landais, Pottock, Shetland, Welsh A and Welsh B).

Personality assessment

Description of the behavioral evaluations. Simplified personality tests were adapted from previous studies (Visser EK et al., 2003; Hausberger et al., 2004; Burger et al., 2007; Lansade et al., 2008c; von Borstel et al., 2011; Munsters et al., 2013; Ijichi et al., 2013; Rothmann et al., 2014). The tests chosen have previously been validated and their stability over time and situation verified (Lansade et al., 2016). The different situations are described

below in the order in which they were conducted and the behaviors recorded are outlined in Tables 1a, 1b and in Table S1 (supplementary data).

(1) CR - height measurement at the withers with a metric stick (HE). The animal was approached by a person with a measuring stick from a distance of 2 m from the left shoulder and the horizontal part of the stick was placed on the withers and the height measured.

(2) CR - loose in the show ring and jumping (JU). The animal was set free in an oval show ring (sand, 18 m x 36 m) 30 seconds (s) before the beginning of a free jumping round. The animal was then also observed during the free jumping round.

(3) CR - conformation assessment (CO). A panel of judges assessed the conformation and gaits of each horse in hand performing a halt (1 minute), walking in one direction and then back, trotting in one direction and then back, followed by a second halt (30 s).

(4) ST - tactile sensitivity (TA). The response to 4 Von Frey type filaments each applied once at the withers (0.008 g on left, 0.02 g on right, 1 g on left and 50 g on right) was assessed (Lansade et al., 2008c; de Sousa et al., 2014).

For safety reasons, before all the fearfulness ST, the handler was given a 4-6m long flexible lunge rein and gloves and was asked to use a slack rein and not to talk to the animal.

(5) ST - fearfulness of a novel object (OB). The animal and its handler began at a starting point 10 m in front of a novel object (1.3m x 1.7m x 0.7m) and then moved around the object which was on their left, trying to keep within a zone of 0 to 2 m from the object before returning to the starting point.

(6) ST- fearfulness of novel surface (SU). A bucket of food pellets was presented under the animal's mouth, 5 m in front of the edge of a novel surface (a 3x3m green tarpaulin with edges maintained by sand). The bucket was then placed in the centre of the novel surface. A maximum of 90 s was given for the handler to move towards the centre of the tarpaulin with the animal on a slack leading rein and for the animal to put its head in the bucket. The handler

initially stayed motionless for 45 s as close as possible to the bucket and then if necessary incited the animal to move with one to three gentle tugs on the leading rein. The reason why food was used here was firstly that it incited the animal to advance onto the tarpaulin and secondly, being placed on the tarpaulin, it encouraged the animal to lower its head when walking forward, to prevent any sudden movement that could have hurt the handler.

(7) ST - fearfulness of suddenness (SD). A trained investigator positioned themselves 5 m from the animal's forelegs. Once the animal was immobile in front of the person, the latter rapidly opened and closed a large black umbrella held on their knees. The test was repeated with the experimenter closer to the animal (3 m). These distances were the same as in Burger et al (2007).

All the ST tests were performed after the CR tests to avoid a possible influence of ST tests on CR tests.

Interpretation of behavioral observations.

We considered that:

- avoidance variables (time to approach and evasion distance) measured during novel surface, novel object and suddenness ST represented fearfulness (Visser et al., 2003; Hausberger et al., 2004; Lansade et al., 2008a).

- the number of whinnies during ST and CR indicated search for social contact (gregariousness) (Waring, 2003; Hausberger et al., 2004; Lansade et al., 2008c)

- the number of steps at the horse initiative or a faster gait could be interpreted as measures of spontaneous locomotor activity (Lansade et al., 2008c) but also as below,

- the number of steps, a more rapid gait, an elevated posture or evasive behaviors during new situations or in a new environment could be interpreted equally as a sign of reactivity/ emotionality/anxiety (Visser et al., 2003; McCall et al., 2006; Momozawa et al., 2007; Rothmann et al., 2014), fearfulness (Hausberger et al., 2004; Lansade et al., 2008a),

nervousness (Hausberger et al., 2004), gregariousness (Hausberger et al., 2004; Lansade et al., 2008c; Rothmann et al., 2014) or stress (Young et al., 2012; Yarnell et al., 2013).

Personality assessors. Multiple assessors (77) recorded the tests. For rating, the assessors and their helper from the French Institute for Horses and Riding had followed a 2-day course on equine ethology and personality tests and 2 days of practice with these tests. All observations were reported on paper with the assessors ticking the correct line corresponding to the precise description of the behavior.

Basic statistics of the personality characteristics. Table 2 (a and b) summarizes the basic statistics of the personality characteristics and the number of animals tested at each stage. Among the 28 characteristics, 19 were responses recorded with scores (1 to 2, 3, 4 or 5) and treated as ordinary responses (Table 2a) and 9 were considered as continuous (Table 2b).

Performance in competition

In France, horses and ponies may participate in official jumping competitions from 4 years of age. Jumping competitions are divided into those for all horses and those reserved for ponies (animals less than 1.50 m and ridden by riders under 18 years of age). From 2014 to 2019, all animals in our study may have participated in official jumping competitions from 4 to 7 years of age or from 4 to 6 years of age for 36 animals born in 2013. Leisure ponies were removed from these analyses.

Of the 876 horses (SH+AA), 724 (83%) took part in at least one jumping competition for horses between 4 and 7 years old and of the 424 sport ponies (SP), 313 (74%) took part in at least one jumping competition for ponies between 4 and 7 years old. Performance indexes in each type of competition were official annual summaries of results in the different competitions (hPI for horses, pPI for ponies). Two criteria were used to calculate the annual hPI and pPI: one was based on the repetition of ranking in each event using maximum

likelihood to estimate the level of the event and the other was based on the sum of points allocated to ranking and the technical difficulty of the event (Ricard et al. 2010). The two PI were corrected for the effect of age, sex and year.

For the French populations of horses and ponies, hPI and pPI were standardized to a mean of 100 and a standard deviation of 20 (except at 4 years old for which standard deviation was lower: 15 for horses and 16 for ponies). For the horses and ponies of our sample, two parameters were used to describe their competition performance: 1) hPI and pPI at 4 years old (hPI4 and pPI4) that measured the particularity of early performance and 2) mean of hPI and mean of pPI from 4 to 7 years old (hPIm and pPIm) that measured the overall aptitude to competition. These means were weighted by $\frac{nr}{1+(n-1)r}$ (with n being the number of years of performance and r the repeatability of the trait, i.e., 0.45 for horse competitions and 0.47 for pony competitions). In our study, the mean of hIP4 was 99.5 (n=519) with a standard deviation of 13.3 and the mean of pIP4 was 101.9 (n=190) with a standard deviation of 18.3. The mean of hPIm was 99.7 (n=724) with a standard deviation of 9.4 and the mean of pPIm was 101.1 (n=313) with a standard deviation of 10.9.

Statistical analyses

Influence of testing conditions and animal related factors on personality characteristics.

The influence of testing conditions and animal related factors were analysed using a general linear model for continuous variables and logistic regression using a cumulative model for ordinary response variables. GLMSELECT and LOGISTIC procedures from the statistical analysis system SAS were used respectively. Effects were tested simultaneously and the best model was retained using stepwise selection with a significance level for entry and staying in the model of 5%. The effects tested were age (3 or 2 years old), sex (male, gelding, female), breed (SH, AA, SP, LP) and breeding show (maximum 67 levels).

Phenotypic relationships between personality characteristics. Phenotypic relationships between personality characteristics were studied using multiple factor analysis (MFA) to balance the weight of each ST or CR in the overall model. First, missing values were imputed with the R package missMDA (Husson and Josse, 2016) and then the MFA analysis was performed using the FactoMineR package (Husson et al., 2016). For the MFA, seven groups were defined which represented the four standard tests and the three customary round assessments.

Phenotypic relationships between personality and competition performance. Regression analysis was performed to evaluate the relationship between each behavioural characteristic and the performance characteristics: hPI4 and hPIm for SH and AA horses, pPI4 and pPIm for sport ponies (SP). In these regressions, the personality characteristics were probit transformed and corrected for significant testing conditions and animal-related factors

Results

Influence of testing conditions and animal-related factors on personality characteristics

The breeding show effect representing the influence of testing conditions was significant for 23 of the 28 characteristics (Table 3). This was the main effect to be taken into account in the model. The characteristics which were less sensitive to the breeding show effect were whinnies and evasive behaviors.

The age effect was significant for nine characteristics (Table 3). Compared to 3-year-old animals, the 2-year-olds were more reactive, moved more and whinnied more than older animals. However, the 2-year-old animals moved at a slower gait when they entered the jumping area (Figure 1a).

During most CR, compared to geldings and females, males whinnied more, moved more and their posture was tenser. However, during all the fearfulness ST geldings reacted more than males or females (Figure 1b).

Few differences were found between breeds except for tactile sensitivity: AA were the most sensitive to the four filaments, followed by SH, then SP and finally LP. The only other breed difference was that LP were more rapid than all the other breeds to put their heads in the bucket on the novel surface (Figure 1c).

Phenotypic relationships between personality characteristics

The first 10 components among the 28 of the MFA analysis were significant (explaining more than 10/28 of the variance). The total percentage of variance explained by these 10 components was 77%. The first factor (16% of the variance) was the combination of posture and evasive behavior characteristics measured in almost all ST and CR and also the fearful measurements during ST (novel object and suddenness, but not novel surface) (Figure 2). The second factor opposed characteristics of novel surface and tactile sensitivity ST in one direction to those of free jumping CR in the opposite direction. The third factor was mainly represented by tactile characteristics. However, characteristics were grouped more by CR or ST than by their type of behavioral measurement such as posture or evasive behavior, except for whinnies. Whinnies measured during the jumping CR and all the ST were grouped along axis 2.

Phenotypic relationships between personality characteristics and jumping performance

Table 4 summarizes the significant relationships between jumping performance in competitions at 4 years old (hPI4, pPI4) and from 4 to 7 years old (hPIm, pPIm) and personality characteristics in horses and in ponies. The few regressors identified were mainly during CR. For horses, these were three characteristics of locomotor activity or reactivity (gait

at entrance, evasive behaviors and posture) during free jumping or conformation CR and one characteristic of gregariousness (whinnies) during free jumping. They were positively linked either to the performance index at 4 years old or to the mean performance from 4 to 7 years old. For sport ponies, only whinnies during height measurement CR and during all the specific tests were significantly positively linked to the mean performance from 4 and 7 years old.

Discussion

Influence of testing conditions and animal-related factors on personality characteristics

The breeding show effect was significant for the majority of personality characteristics. This effect included variation of environmental conditions (place, temperature, hour, wind, noise, effect of other horses or ponies) and assessment (different assessors rating personality), without being able to distinguish between these two effects. To minimize this second effect and optimize their reproducibility, the situations for performing ST were described in detail and easy to set up, the measurements during CR and ST were very simple to carry out, and the assessors were trained and used observation sheets that were informative and easy to fill out. An improved estimation of the breeding show effect would be obtained if the number of animals assessed at the same place was high and if the assessor effect could be calculated and removed. This could be possible if the same assessors rated personality at several breeding shows.

The effect of age on personality characteristics was demonstrated in our study: 2-year-old animals were more fearful and reactive and more gregarious than 3-yearolds. This is in line with previous findings for fearfulness and reactivity for young horses (9 vs. 22 months; (Visser et al., 2001)) and for older horses (3 to 24 years; (Baragli et al., 2014; Graf et al., 2014)).

In the current study, geldings were found to be slightly more fearful than females or males only during the fear ST. In the literature, findings on sex effects on equine fearfulness and reactivity are contradictory. Some studies found no difference between geldings and females during fear tests (Hausberger et al., 2004; Lesimple et al., 2011; Baragli et al., 2014; Graf et al., 2014). While another study reported very slight differences between stallions and mares during a flashlight test and during transportation (Ishizaka et al., 2017). Other research indicated that young males showed a higher cardiac frequency than young females during novel situations (Janczarek and Kędzierski, 2011), and young geldings exhibited a higher reactivity than young females in a fear test in an open field (Lansade, unpublished data). On the other hand, stallions were observed to be less fearful than geldings or mares (Graf et al., 2014) and young males were found to be less fearful than young females (Oki et al., 2007). However, in some of this literature, there were confounding factors like very different breeds and ages (Hausberger et al., 2004; Lesimple et al., 2011; Graf et al., 2014) even though there were corrections or an unspecified type of male (intact or castrated) (Oki et al., 2007; Janczarek and Kędzierski, 2011). Therefore, the question about the influence of sex on fearfulness remains open.

However, in situations other than fear tests, in our study, males whinnied more, moved more and had a tenser posture than geldings or females. This result agrees with other studies on horse behaviour. Stallions tend to adopt more rapid movement than females under natural conditions (Duncan, 1980). They obtained higher notes in movement evaluation at walk in breeding shows (Wejer and Lewczuk, 2016). Moreover, movements may accompany whinnying (Waring, 2003). They could be seeking social contact because they are often housed alone (Irrgang and Gerken, 2010). They could be differently affected by sight or smell of other horses including mares (Guillaume et al., 2018). All these facts could explain the

disparity of the sex effect found in the literature and which complicate the interpretation of personality characteristics.

Personality differences between breeds observed in this study concerned tactile sensitivity, with a lower sensitivity measured in ponies than in horses, particularly Anglo-Arabians. We have already observed lower tactile sensitivity in Merens stallions (a mountain breed) than in French jumping stallions (Vidament et al., 2015). Thus tactile sensitivity is certainly higher in light breeds than in draught or pony breeds. Fearfulness was the other trait that distinguished the leisure ponies from the other breeds, but only during one fear ST. These fearfulness or reactivity differences between breeds (sport horses being more fearful than ponies or draught horses) have been highlighted by numerous authors (Hausberger et al., 2004; Lesimple et al., 2011; von Borstel et al., 2012; Graf et al., 2014).

In conclusion, concerning the influence of the testing conditions and animal-related factors on the personality assessment, the results obtained in the current study with this very simple protocol confirmed most of the earlier findings. Therefore, this protocol provides relevant behavioral information. With respect to phenotypic information, attention should be paid to how these fixed effects are taken into account in the model.

Phenotypic relationships between personality and jumping performance

Performance underlies many other factors than personality alone, including the animal physical characteristics, its training, the rider and environment, thus the relationship between personality and performance could be expected to be weak. In our study, in horses, among the four different personality characteristics linked to performance, three were from the jumping CR (whinnies, gait on entering and evasive behaviors) and one from the conformation CR (posture). In ponies, only whinnies in two situations were linked to performance. Although

329 significant, these links were weak and only one characteristic was actually linked to the
330 performance both in horses and in ponies: whinnies during the different CR or ST.

331 The literature on the relationship between personality and jumping abilities is limited. In a
332 study by Visser et al. (2003), personality tests measuring emotionality were performed at 1
333 and 2 years old on about 30 horses. Then, at 3 years old, horses were trained over jumps and
334 scored for their jumping capacity during this training and over a novel course of obstacles
335 (correct jumps and refusals). Authors found contradictory correlations with some personality
336 characteristics measured at 1 and 2 years old, for example a positive correlation between the
337 latency to touch a novel object and performance at 1 year old and a negative correlation at 2
338 years old. In a study by Lansade et al. (2016), two groups of 15 to 24-horses were evaluated
339 with simplified personality tests either at 3 or at 4-5 years old and then their performance
340 during their second show jumping competition of the year at 4 years old for the younger
341 group or during the first five show jumping competitions of the year for the older group was
342 scored (number of penalties, refusals, veering from the center before jumping). Several
343 positive associations between fearfulness and performance were found. Using a larger sample
344 of 3-year-old Danish warmblood (n = 234), in field conditions similar to the ones used in our
345 study, Rothmann et al. (2014) found negative correlations between reactivity (a score based
346 on body and head movements, snorting and defecations) during conformation evaluation and
347 free jumping ability (-0.14) or rideability (-0.16) measured once, on the same day.
348 Correlations found by Rothmann et al. (2014) were low, thus consistent with the quasi-
349 absence of correlation found here.

350 Moreover, the only common characteristic found to be related to jumping performance in
351 the two populations (horses, ponies) was the whinnies emitted in different tests. They are
352 indicative of search for social contact and are considered as a good behavioral parameter to
353 evaluate the dimension of gregariousness during a social isolation test (Waring, 2003;

354 Hausberger et al., 2004; Lansade et al., 2008b). Intact males whinny more than geldings or
355 mares (see paragraph above). In general, the best jumping males are kept uncastrated to keep
356 them for reproduction, even though they may be difficult to handle. This raises the question of
357 whether the relationship found between whinnies and performance could be due to the
358 presence of intact males in our two populations. However, in our study, this behavior
359 characteristic which had a low variability (only 11 to 23 % horses or ponies whinnied during
360 the different tests) was corrected for sex (female, gelding, intact male) and the official
361 performance data were corrected for sex (female and male (intact or not)). This sex effect
362 should therefore not have influenced the evaluation of the relationship between personality
363 and performance. However, grouping intact and castrated males in the same sex category in
364 the performance index may have introduced a bias. In conclusion, this observation on
365 whinnies is rather difficult to interpret.

366 In our study, two of the four personality characteristics related to jumping performance in
367 the population of horses are locomotor/reactivity characteristics measured during the free
368 jumping CR. This CR is a preparatory step for jumping competitions. This result is in line
369 with Lansade (2016) who found a link between the “main gait during the first 30 s before free
370 jumping” characteristic and success in the first five show-jumping competitions of the year. It
371 would concern the horses that had the most experience and showed their anticipation of
372 entering the free jumping ring. To support this hypothesis, among the characteristics that were
373 different between 2- and 3-year-old animals in our study, the main gait during the first 30 s
374 was the only variable that was higher at 3 years old than at 2 years old, while all the others
375 were lower. So this relationship could come from learning and anticipating.

376 Our study also provides information about the tactile sensitivity characteristics, little
377 studied in horses until now. Another way of presenting the results would be to evaluate the
378 percentage of animals responding to each weight of filament. The horses and ponies in our

379 study responded as follows: 0.008 g (18% of the population), 0.02 g (29%), 1 g (78%) and 50
380 g (72%). We thus confirmed earlier findings. First, animals that respond to a filament of
381 between 0.5 and 10 g did not necessarily respond to higher pressures which would be quite
382 different from those of insects on the skin (Lansade et al., 2008c). Second, tactile sensitivity is
383 independent of fear characteristics (Lansade et al., 2008c; Lansade et al., 2016). Third, no
384 relation between tactile sensitivity and jumping performance is coherent with the very weak
385 link found by Lansade et al. (2016). However, in horses, variation in tactile sensitivity has
386 also been associated with mood disorders (hyposensitivity in depression-like syndrome
387 (Fureix et al., 2012) or with stereotypy (hypersensitivity in crib-biting horses (Freymond et
388 al., 2019)).

389 Taken together, all these studies suggest that there is no, or at best, only a weak
390 relationship between jumping performance and personality characteristics. As no main fear or
391 tactile sensitivity variables were significantly related to jumping performances in the two
392 populations studied, it seems that less fearful animals or animals with less tactile sensitivity
393 may perform well in show jumping.

394 *Breeding and training considerations*

395 Fearful horses are considered inappropriate for amateur riders for safety reasons and their
396 lower rideability (Visser et al., 2008; Rothmann et al., 2014; Lansade et al., 2016). For
397 behavioral tests to be included in the selection process it is necessary to estimate their
398 heritability and their genetic correlation with the traits that express breeding goals. However,
399 in our study the size of each population (horses and ponies) was insufficient to reliably
400 estimate these correlations in each population. Nevertheless, research has shown that in
401 French sport horses, the genetic correlation between performance in amateur show jumping
402 horses (fence height: <=1.25 m) and performance in professional show jumping horses

403 (>=1.30 m) was very high (0.95) (Ricard, 2016). This was calculated on 70,590 amateur show
404 jumping horses and 16,028 professional show jumping horses. Among them, 15,409 had
405 results in both types of competitions. These horses originated from 257,427 horses (6
406 generations). This therefore indicates that jumping traits for French amateur and professional
407 show jumping horses are genetically very close.

408 Habituation to stress situations is part of horse breeding and training. Some behavioural
409 tests particularly fear tests run counter generally recommended training methods (McLean and
410 Christensen, 2017). To our knowledge, no research has investigated the effect of fear tests on
411 the relationship between horses and their trainers when horses are in hand. However, some
412 studies have described in detail the physiological (autonomic and endocrine change) and
413 behavioral effects involved in a suddenness test (McCall et al., 2006; Lansade et al., 2008a;
414 Villas-Boas et al., 2016; Scopa et al., 2018). In real life, horses experience many fear-
415 inducing situations: for example, flapping flags, running/barking dogs , first vaccination, first
416 trimming, first loading onto a trailer, or first time in an automated horse walker (Janczarek
417 and Kędzierski, 2011). Consequently, habituation to fear stimuli often plays an integral part of
418 training of young horses. However, this habituation should be conducted gradually
419 (Christensen et al., 2006) to avoid animals becoming sensitized. In our study, the animals
420 were tested in only one breeding show and should not have previously received any specific
421 training for these tests although this could not be guaranteed. Each fear test was very short
422 and only 6 minutes on average were required to complete all three tests. Thus the immediate
423 side effects of these fear tests should be minimal.

424 *Potential risks of loss of quality in the personality data collection*

425 Certain factors limit the degree to which our results can be interpreted and generalized.
426 Firstly, animal behavior was evaluated directly without video. For practical reasons, it was not

427 possible to film all the CR and ST. Instead we chose to train the assessors in observation
428 techniques and use detailed rating grids. For further application by the breeders' associations,
429 it was important to enable data to be collected under real-life conditions

430 The fact that the animals were held in hand by their owner or trainer could be considered
431 as a possible bias for all CR and ST observations except during free jumping. Indeed, horses
432 tend to have lower reactions when held in hand than when free or ridden (Górecka et al.,
433 2007; von Borstel et al., 2011). We chose to have horses in hand as it represented the main
434 condition during breeding shows, it was also safer for the animals and for reasons of liability.

435 Another limitation was that the animals involved in this study certainly had varying levels
436 of handling and training, but this is always the case in breeding shows and competitions. The
437 objective was to assess young horses which had only limited experience. In the event of a
438 horse being tested in a second breeding competition, the data from the second competition
439 was removed from the database. As this study represented one of the first times the different
440 CR or ST situations had been used, owner/ trainer knowledge of them was limited and prior
441 training for them unlikely. In addition, the results of the personality tests did not impact the
442 other results of the breeding show.

443 It would be interesting to evaluate the repeatability of these tests, but this would require
444 repeating the same tests on the same horses under similar conditions. This was not possible
445 under our field conditions and when by chance it occurred we kept only data of the first
446 breeding show. Moreover, the limitation mentioned above would persist, namely the risk that
447 the horses might be trained differently between the two evaluations. Von Borstel et al. 2012
448 found that the repeatability of scores obtained by ridden horses during three fear tests were
449 similar and high (around 0.7). Stability of fear reactions and tactile sensitivity over time and
450 situations has already been demonstrated for our tests (Lansade et al., 2008a, 2008c, 2016).
451 Von Borstel et al. 2012 proposed assessing the decrease in reactivity between two repeated

tests, preventing any interest in deliberately training horses for the first evaluation. However, this would double the time required and thus would only be possible for valuable horses and for tests eliciting moderate fear."

The participation rate of animals in the CR and ST personality assessment in a given breeding show should be noted. The personality observations for CR were conducted on 100% of animals, but for ST this rate was 50-70% because owners had the choice to participate or not. This implies a potential bias, as animals presenting problematic behaviors could have been withheld from participating in the ST. However, as previously stated, the ST were performed after the CR to avoid influencing the CR and the results of the personality evaluation were not known to the judges who ranked the animals at the end of the breeding show on morphology and gaits recorded during the event. Moreover, there was no indication of selection bias, based on earlier pre-selection process for breeding shows, which would have influenced the correlation with jumping performance. The studied population was representative for the French population as the mean performance indexes between 4 and 7 years old of each sub-population tested (99.7 for the horses and 101.1 for the ponies) were very similar to that of the whole population of such animals in France (mean of these indexes: 100 by construction). Finally, the decrease in the number of animals assessed for personality between 2013 and 2015 may be a limitation. This decrease was due to both the number of breeding shows and of animals participating per show falling considerably in France over that period. Nevertheless, the participation rate in ST remained relatively stable at 50 to 70% when these tests were proposed, so this decrease can be eliminated as a bias in the collection of personality data.

Conclusions

This study involving more than 800 horses and 400 ponies assessed for personality at an early age and then ridden in official competitions (for 80% of them) showed that 1) fixed effects of breeding show, age, sex and breed were significant in the analysis model of personality characteristics measured at 2 or 3 years old, 2) jumping performance at 4, 5, 6 and 7 years old was seldom phenotypically correlated with these personality characteristics in the two populations of jumping horses and jumping ponies. Links were evidenced with a very small number of gregariousness and locomotor/reactivity characteristics measured during customary rounds, but not with reactivity assessed during specific fear tests or with tactile sensitivity. Therefore, it seems that less fearful horses and ponies may perform well in show jumping.

Policy and ethics

This study complied with French laws relating to animal experimentation and the European directive 2010/63/EU for animal experiments. The horses tested in this experiment were not research animals. Their husbandry and care were their owners' responsibility. The owners were free to choose to participate or not in the specific tests (ST).

Acknowledgements

The authors thank all the personnel from local breeding societies, particularly the AECVL (horse breeding society of the Val de Loire), and all the IFCE teams who conducted the tests and rated the horses and ponies in 2013, 2014 and 2015 (people from Amboise, Chazey, Compiègne-Amiens, Cluny, Les Bréviaires, La Roche/Yon, Le Lion d'Angers, Pompadour, Rosières, Saintes, Uzès). They also thank S. Edrich, C. Leterrier and S. Love for assistance in preparing the manuscript. This project received funding from the Fonds Eperon (2013).

Conflict of interest statement

The authors declare no conflict of interest. They have no financial or personal relationships that could inappropriately influence or bias the content of the article.

Authorship statement

The idea for the paper was conceived by MV and AR. The behavioural tests and observations were designed by MV, LL and SD. The measurements were organized by MV, BDSP and SD, and then by the teams. The teaching and training of the teams was conducted by MV and BDSP. The data were prepared by MV, MS and AR and analyzed by AR. The paper was written by AR and MV and approved by all.

References

Baragli, P., Vitale, V., Banti, L., Sighieri, C., 2014. Effect of aging on behavioural and physiological responses to a stressful stimulus in horses (*Equus caballus*). Behaviour. 151, 1513-1533.

Bates, J. E. 1989. Concepts and Measures of Temperament. In: Kohnstamm G. A., Bates J. E. and Rothbart M. K., (Ed.), Temperament in Childhood. Wiley, New York. p. 3-26.

Burger, D., Imboden, I., Jallon, L., Ionita, J. C., Rapin, V., Doherr, M., Poncet, P. A., 2007. Introduction of a behavioural test for Franches-Montagnes horses. In: Hausberger, M., Sondergaard, E., Martin-Rosset, W (Eds.), Horse Behaviour and Welfare. EAAP Publication 122, Wageningen, NL, pp. 13-22.

Burrow, H. M., Corbet, N. J., 2000. Genetic and environmental factors affecting temperament of zebu and zebu-derived beef cattle grazed at pasture in the tropics. Aust. J. Agr. Res. 51, 155-162.

Canario, L., Mignon-Grasteau, S., Dupont-Nivet, M., Phocas, F., 2013. Genetics of behavioural adaptation of livestock to farming conditions. Animal. 7, 357-377.

Christensen, J. W., Rundgren, M., Olsson, K., 2006. Training methods for horses: habituation to a frightening stimulus. Equine Vet. J. 38, 439-443.

de Sousa, M. V. P., Ferraresi, C., de Magalhaes, A. C., Yoshimura, E. M., Hamblin, M. R., 2014. Building, testing and validating a set of home-made von Frey filaments: A precise, accurate and cost effective alternative for nociception assessment. J. Neurosci. Meth. 232, 1-5.

Duncan, P., 1980. Time-Budgets of Camargue Horses .2. Time-budgets of adult horses and weaned sub-adults. Behaviour. 72, 26-49.

Feaver, J., Mendl, M., Bateson, P., 1986. A method for rating the individual distinctiveness of domestic cats. Anim. Behav. 34, 1016-1025.

Freymond, S. B., Bardou, D., Beuret, S., Bachmann, I., Zuberbuhler, K., Briefer, E. F., 2019. Elevated sensitivity to tactile stimuli in stereotypic horses. Front. Vet. Sci. 6.

Fureix, C., Jego, P., Henry, S., Lansade, L., Hausberger, M., 2012. Towards an ethological animal model of depression? A study on horses. Plos One. 7.

Górecka, A., Bakuniak, M., Chruszezewski, M. H., Jezierski, T. A., 2007. A note on the habituation to novelty in horses: handler effect. Anim. Sci. Pap. Rep. 25, 143-152.

Graf, P., von Borstel, U. K., Gauly, M., 2014. Practical considerations regarding the implementation of a temperament test into horse performance tests: Results of a large-scale test run. J. Vet. Behav. 9, 329-340.

542 Guillaume, D., Moussu, C., de Geoffroy, F., Chesneau, D., Keller, M., 2018. Olfactory stimulation or inhibition
543 of sexual behavior of stallions in non-breeding season. Physiol. Behav. 186, 1-9.
544 Hall, C., Randle, H., Pearson, G., Preshaw, L., Waran, N., 2018. Assessing equine emotional state. App. Anim.
545 Behav. Sci. 205, 183-193.
546 Haskell, M. J., Simm, G., Turner, S. P., 2014. Genetic selection for temperament traits in dairy and beef cattle.
547 Front. Genet. 5,
548 Hausberger, M., Bruderer, C., Le Scolan, N., Pierre, J.-S., 2004. Interplay between environmental and genetic
549 factors in temperament/personality traits in horses (*Equus caballus*). J. Comp. Psychol. 118, 434-446.
550 Husson, F., Josse, J. 2016. Handling Missing Values with Multivariate Data Analysis, Package 'missMDA',
551 Repository CRAN.
552 Husson, F., Josse, J., Le, S., Mazet, J. 2016. Multivariate Exploratory Data Analysis and Data Mining, Package
553 'FactoMineR', Repository CRAN.
554 Ijichi, C., Collins, L. M., Creighton, E., Elwood, R. W., 2013. Harnessing the power of personality assessment:
555 subjective assessment predicts behaviour in horses. Behav. Processes. 96, 47-52.
556 Irrgang, N., Gerken, M., 2010. An investigation of housing conditions, applied management, handling practises
557 and behaviour in purebred Arabian stallions. Zuchtungskunde. 82, 292-302.
558 Ishizaka, S., Aurich, J. E., Ille, N., Aurich, C., Nagel, C., 2017. Acute physiological stress response of horses to
559 different potential short-term stressors. Journal of Equine Veterinary Science. 54, 81-86.
560 Janczarek, I., Kędzierski, W., 2011. Emotional response to novelty and to expectation of novelty in young race
561 horses. J. Equine Vet. Sci. 31, 549-554.
562 Lansade, L., Bouissou, M. F., Erhard, H. W., 2008a. Fearfulness in horses: A temperament trait stable across
563 time and situations. App. Anim. Behav. Sci. 115, 182-200.
564 Lansade, L., Bouissou, M. F., Erhard, H. W., 2008b. Reactivity to isolation and association with conspecifics: A
565 temperament trait stable across time and situations. App. Anim. Behav. Sci. 109, 355-373.
566 Lansade, L., Philippon, P., Herve, L., Vidament, M., 2016. Development of personality tests to use in the field,
567 stable over time and across situations, and linked to horses' show jumping performance. App. Anim.
568 Behav. Sci. 176, 43-51.
569 Lansade, L., Pichard, G., Leconte, M., 2008c. Sensory sensitivities: Components of a horse's temperament
570 dimension. App. Anim. Behav. Sci. 114, 534-553.
571 Lesimple, C., Fureix, C., LeScolan, N., Richard-Yris, M. A., Hausberger, M., 2011. Housing conditions and
572 breed are associated with emotionality and cognitive abilities in riding school horses. App. Anim.
573 Behav. Sci. 129, 92-99.
574 McCall, C. A., Hall, S., McElhenney, W. H., Cummins, K. A., 2006. Evaluation and comparison of four
575 methods of ranking horses based on reactivity. App. Anim. Behav. Sci. 96, 115-127.
576 McLean, A. N., Christensen, J. W., 2017. The application of learning theory in horse training. App. Anim.
577 Behav. Sci. 190, 18-27.
578 Momozawa, Y., Terada, M., Sato, F., Kikusui, T., Takeuchi, Y., Kusunose, R., Mori, Y., 2007. Assessing equine
579 anxiety-related parameters using an isolation test in combination with a questionnaire survey. J. Vet.
580 Med. Sci. 69, 945-950.
581 Munsters, C. C., Visser, E. K., van den Broek, J., Sloet van Oldruitenborgh-Oosterbaan, M. M., 2013.
582 Physiological and behavioral responses of horses during police training. Animal. 7, 822-827.
583 Oki, H., Kusunose, R., Nakaoka, H., Nishiura, A., Miyake, T., Sasaki, Y., 2007. Estimation of heritability and
584 genetic correlation for behavioural responses by Gibbs sampling in the Thoroughbred racehorse. J.
585 Anim. Breeding Genet. 124, 185-191.
586 Ricard, A. 2016. Elite horse, horse for all: genetically very compatible objectives for show jumping (Cheval
587 d'élite, tous à cheval : des objectifs génétiquement très compatibles pour le saut d'obstacle). In: 42
588 Journée de la Recherche Equine, Institut Français du Cheval et de l'Equitation (IFCE), Paris, France,
589 23-27.
590 Rothmann, J., Christensen, O. F., Sondergaard, E., Ladewig, J., 2014. Behavior observation during conformation
591 evaluation at a field test for Danish Warmblood horses and associations with rideability and
592 performance traits. J. Equine Vet. Sci. 34, 288-293.
593 Scopa, C., Palagi, E., Sighieri, C., Baragli, P., 2018. Physiological outcomes of calming behaviors support the
594 resilience hypothesis in horses. Sci. Rep. 8, 17501.
595 Vidament, M., Yvon, J. M., Le Bon, M., Dumont Saint Priest, B., Danvy, S., L., L. 2015. Temperament of
596 horses measured by standardized tests: relation with age, breed and level of riders (Le tempérament des
597 chevaux mesuré par des tests standardisés : relation avec l'âge, la race et le niveau du cavalier). In: 41
598 Journée de la Recherche Equine, Institut Français du Cheval et de l'Equitation (IFCE), Paris, France,
599 15-24.

600 Villas-Boas, J. D., Dias, D. P. M., Trigo, P. I., Almeida, N. A. D., de Almeida, F. Q., de Medeiros, M. A., 2016.
601 Behavioural, endocrine and cardiac autonomic responses to a model of startle in horses. App. Anim.
602 Behav. Sci. 174, 76-82.
603 Visser, E., Van Reenen, C., Engel, B., Schilder, M., Barneveld, A., Blokhuis, H., 2003. The association between
604 performance in show-jumping and personality traits earlier in life. App. Anim. Behav. Sci. 82, 279-295.
605 Visser, E. K., Van Reenen, C. G., Blokhuis, M. Z., Morgan, E. K. M., Hassmen, P., Rundgren, T. M. M.,
606 Blokhuis, H. J., 2008. Does horse temperament influence horse-rider cooperation? J. App. Anim. Welf.
607 Sci. 11, 267-284.
608 Visser, E. K., van Reenen, C. G., Hopster, H., Schilder, M. B. H., Knaap, J. H., Barneveld, A., Blokhuis, H. J.,
609 2001. Quantifying aspects of young horses' temperament: consistency of behavioural variables. App.
610 Anim. Behav. Sci. 74, 241-258.
611 von Borstel, U. K., 2013. Assessing and influencing personality for improvement of animal welfare: a review of
612 equine studies. CAB Reviews. 8, 1-27.
613 von Borstel, U. K., Euent, S., Graf, P., Konig, S., Gauly, M., 2011. Equine behaviour and heart rate in
614 temperament tests with or without rider or handler. Physiology & Behavior. 104, 454-463.
615 von Borstel, U. K., Pirsich, W., Gauly, M., Bruns, E., 2012. Repeatability and reliability of scores from ridden
616 temperament tests conducted during performance tests. App. Anim. Behav. Sci. 139, 251-263.
617 Waring, G. H., 2003. Horse Behavior. 2nd ed. Noyes Publications/William Andrew Publishing, Norwich, NY,
618 USA.
619 Wejer, J., Lewczuk, D., 2016. Effect of the age on the evaluation of horse conformation and movement. Ann
620 Anim Sci. 16, 863-870.
621 Wolyncewicz, G. E. L., Palmer, C. S., Jowett, H. E., Hutson, J. M., King, S. K., Teague, W. J., 2018. Horse-
622 related injuries in children - unmounted injuries are more severe: a retrospective review. Injury. 49,
623 933-938.
624 Yarnell, K., Hall, C., Billett, E., 2013. An assessment of the aversive nature of an animal management procedure
625 (clipping) using behavioral and physiological measures. Physiol. Behav. 118, 32-39.
626 Young, T., Creighton, E., Smith, T., Hosie, C., 2012. A novel scale of behavioural indicators of stress for use
627 with domestic horses. App. Anim. Behav. Sci. 140, 33-43.
628
629

Figure captions

Figure 1. Fixed effects of sex, age, and breed on behavioral characteristics during personality evaluation of young horses and ponies during customary rounds (CR) and specific tests (ST) in breeding shows. Only significant effects are shown ($P<0.05$). The results are presented after transformation of variables with standard phenotypic deviation set to 1. The most frequent category (3 years for age, female for sex, Sport Horses for breed) is the reference, its mean is set to 0. The colored bars are the differences between this reference and the other categories. The black bars represent the standard errors of these differences.

a. Age effect (2 years vs 3 years = reference) in standard phenotypic deviation unit. The 2 years population is compared to the 3 years population whose mean is set to 0 and which is used as the reference population.
b. b. Sex effect in standard phenotypic deviation unit (Gelding, Male vs Female (= reference)). The Geldings (castrated males) and Males (non castrated males) populations are compared to the Females population whose mean is set to 0 and which is used as the reference population.
c. c. Breed effect in standard phenotypic deviation unit (AA=Anglo-Arabian, LP=leisure ponies, SP=sport ponies vs Sport Horses (SH) (=reference)). The AA, LP, SP populations are compared to the SH population whose mean is set to 0 and which is used as the reference population.

Meaning of abbreviations:

During height measurement CR, HE_tim: time to measure the horse; HE_ste: number of steps, HE_pos: posture, HE_whi: whinnies

During jumping CR, JU_gait: gait, JU_pos: posture, JU_eva : evasive behaviors, JU_whi :whinnies,

During conformation CR, CO_ste: steps during two halts, CO_pos, posture, CO_eva: evasive behaviors, CO_whi: whinnies.

During tactile sensitivity ST , TA_num: number of filaments leading to quivering responses, TA_min: minimal strength of filament leading to a quivering response.

During novel object ST, OB_dis: mean of evasion distances, OB_tim: time to turn around the object.

During novel surface ST, SU_foo: latency to put first feet on the surface, SU_hea: latency to put the head in the bucket on the surface, SU_pos: posture.

During suddenness ST, SD_fli: mean flight.

During all the ST, ST_whi: whinnies

Figure 2. First two components of Multiple Factor Analysis of behavioral characteristics measured during personality evaluation of young horses and ponies during customary rounds (CR) and specific tests (ST) in breeding shows.

Meaning of abbreviations:

During height measurement CR (dark green), HE_tim: time to measure the horse, HE_ste: number of steps during two halts, HE_pos: posture, HE_eva: evasive behaviors, HE_whi: whinnies

During jumping CR (light blue), JU_gait: main gait during the first 30 s before jumping, JU_pos: posture, JU_eva : evasive behaviors, JU_whi :whinnies, JU_def: defecation

During conformation CR (light green), CO_ste: mean of steps during two halts, CO_pos, posture, CO_eva: evasive behaviors, CO_whi: whinnies.

During tactile sensitivity ST (purple), TA_num: number of filaments leading to quivering responses, TA_min: minimal strength of filament leading to a quivering response.

During novel object ST (brown), OB_dis: mean evasion distance from the object, OB_tim: time to turn around the object, OB_pos: posture, OB_eva: evasive behaviors

During novel surface ST (pink), SU_bef: latency to eat before the test, SU_foo: latency to put first feet on the surface, SU_hea: latency to put the head in the bucket on the surface, SU_pos: posture, SU_eva: evasive behaviors.

During suddenness ST (light brown), SD_dis: mean evasion distances, SD_fli: mean flight movement scores.

During all the ST (light brown), ST_whi: whinnies

Figure 1a : Age effect (2 years vs 3 years = reference)

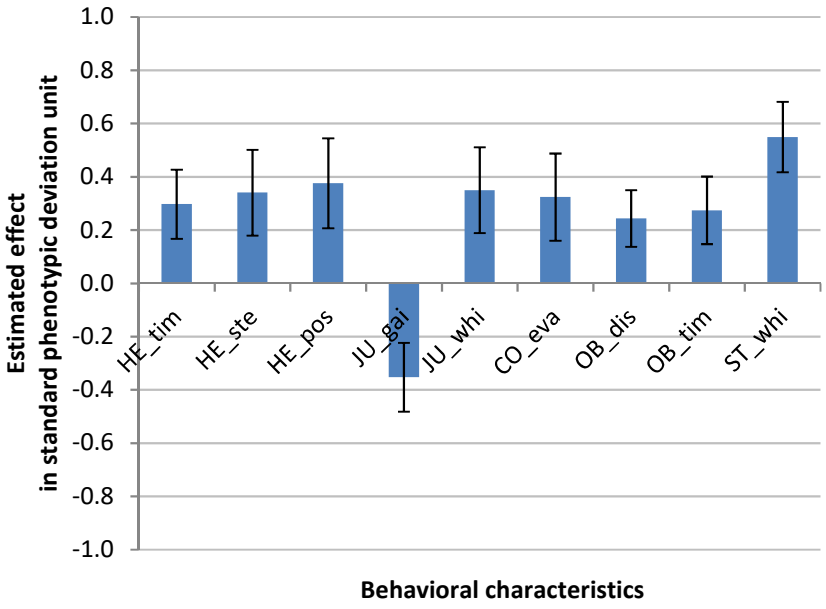


Figure 1b : Sex effect (Gelding, Male vs Female (= reference))

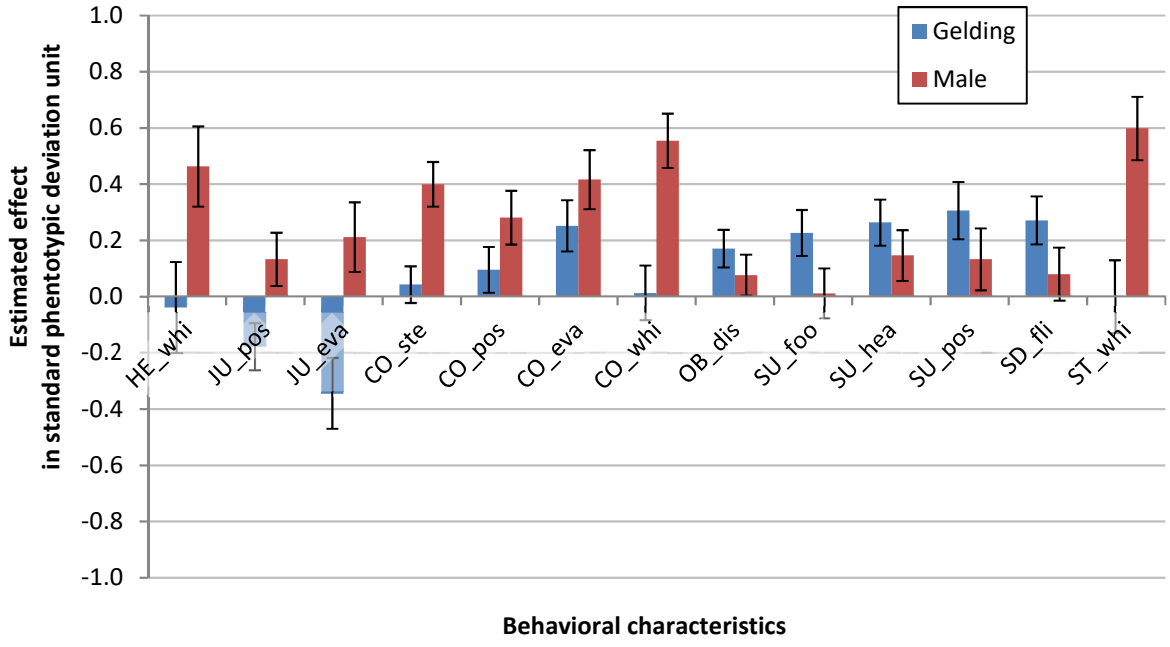


Figure 1c : Breed effect (AA, LP, SP populations vs SH population = reference)

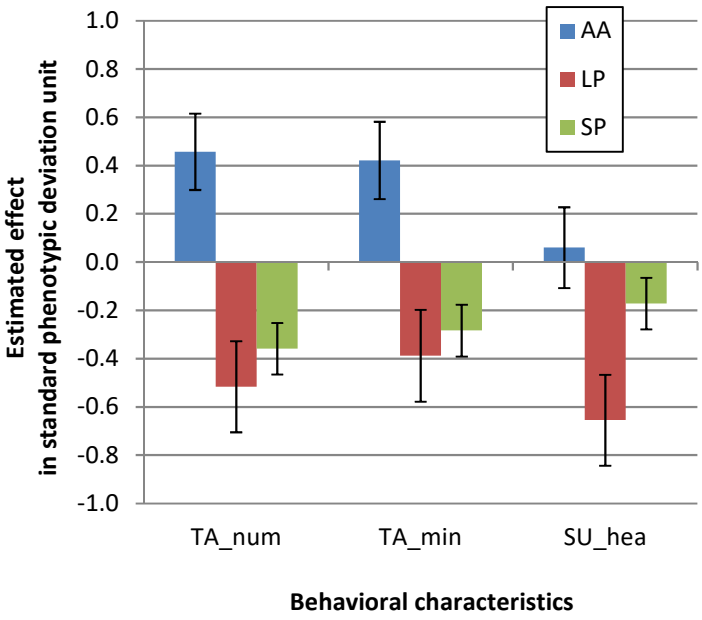
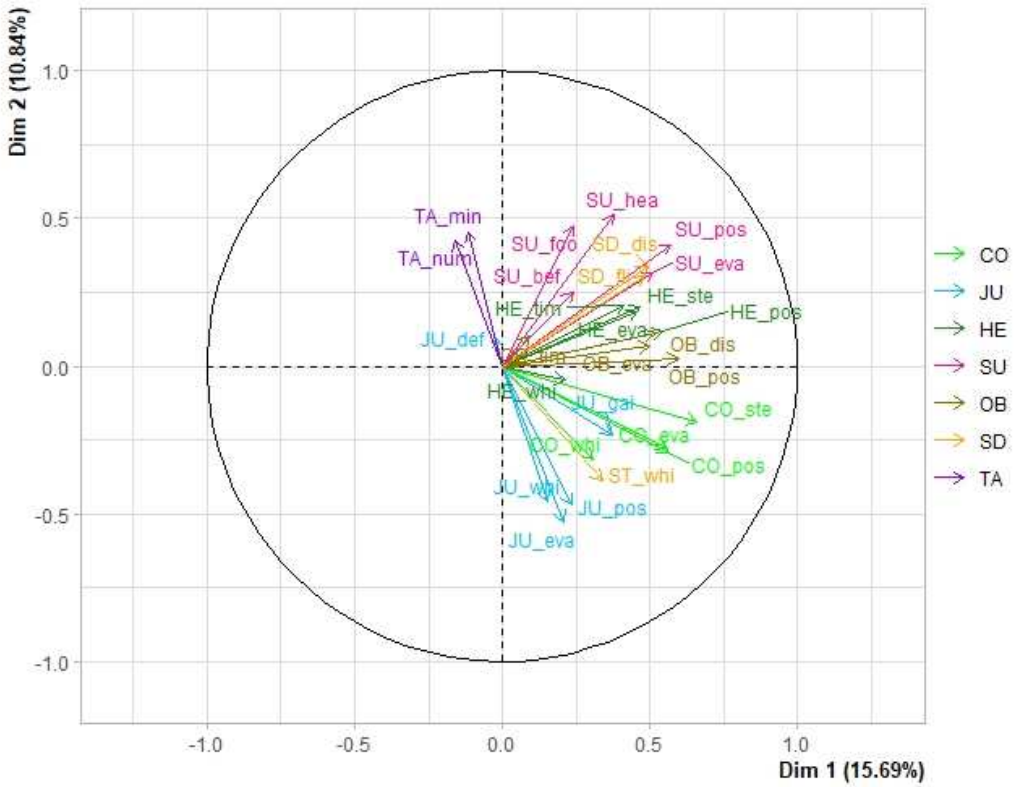


Figure 2.


Dim 2 (10.84%)
1.0
0.5
0.0
-0.5
-1.0
-1.0
-0.5
0.0
0.5
1.0
Dim 1 (15.69%)
TA_min
TA_num
SU_hea
SU_foo
SD_dis
SU_pos
SU_eva
SU_bef
SD_fla
HE_ste
HE_pos
HE_eva
JU_def
OB_dis
OB_pos
HE_whi
JU_cal
CO_ste
CO_eva
CO_whi
CO_pos
ST_whi
JU_whi
JU_pos
JU_eva
CO
JU
HE
SU
OB
SD
TA