

Chapter 7

Discussion

The aim of this chapter is to discuss the overall aim of the dissertation which was to determine a set of objective behaviour tests for the selection of horses with a personality suitable to train as police horses. This aim was achieved, and the main finding from the validity assessment of the standardised behavioural tests in predicting suitability was some evidence of association of horses' suitability for a particular role and response to reactivity tests.

These associations were between a horses' postural expression to an unexpected noise and its suitability at community events, the horses' flight reaction to an unexpected noise and its suitability at community events and for training officers, the horses' flight reaction to a sudden moving object and its suitability for local patrols, public order and for training officers and the horses' tendency to touch a sudden moving object and its suitability at community events and training officers. An association was also found between the horses' latency to touch a stranger and its suitability at community events.

The other behaviour tests, as in other studies, were found to have poor correlations in responses to reliable standardised behaviour tests in predicting the suitability of adult horses (as a police horse)/riding horses. Le Scolan et al. (1997), found only moderate correlations between riding school horses' reactions to an unusual surface and subjective ratings, they also

failed to establish the reliability of their tests. Seaman et al. (2002) and Momozowa et al. (2003) also found behaviour tests to be invalid in predicting horses' general behavioural tendencies, such as their reaction to being groomed and handled. This is contrary to Visser et al. (2003) results, which found correlations between the horses' responses to standardised behaviour tests and suitability scores, however the horses used in their study were youngsters.

Limitations of the validity assessment

The first limitation was the restricted number of horses which could be used at the police units. In the planning of the study, it had been arranged for more horses to be used (35), but due to health, behavioural and time restraints the sample size was decreased to 24 horses. The fast moving object test also had to be excluded from one of the police units as they were not comfortable with letting the horses loose in the arena when opening the umbrella, this lowered the sample size down to 15 horses for this test.

This limitation will have affected the results, as a larger sample size would have produced more conclusive evidence as the larger the sample size, the better the probability that the sample statistics will closely approximate the population values (Lehner, 1996). However in behavioural science this situation is often unavoidable, due to limited numbers of subjects (Martin & Bateson, 1993) and for the majority of tests where $N=24$ the sample size was sufficient for the results to stand. However for the umbrella test $N=15$, the sample was possibly too small for the results to be conclusive.

The second limitation prevented the researcher from handling the horses directly, meaning the experimenter could not be kept constant at both units. In Hasberger and Muller's (2002) study it was found that horses reacted differently to being approached, depending on the caretaker. Although the testers were instructed on exactly how to conduct the tests, the way they handled the horses may have had an effect on the horses' behaviour, although this would not have influenced their reactions in the reactivity tests. The testers were also not completely unknown to the horses, as although they did not work with the horses on a daily basis, they were familiar with horses and may have had pre expectations about how the horses would react.

Secondary analysis (see Table 6.5.1.1) was used to explore whether there was a difference in horses' behavioural reactions depending on their location, the results indicated no statistically significant differences between the locations after Bonferroni correction, indicating the police horses' behaved similarly at both units, and so were not influenced to any great degree by the identity of the handler.

The third limitation is concerned with the time of year the tests were conducted (between January and March), the horses were being stabled for long periods of time so were likely to have been quite lively. The tests were also conducted over weekends when the horses were not in work and so possibly had more energy, causing them to be more reactive. This could have had potential effects on the results as horses may have been more

reactive than normal during the data collection. The horses are normally only ridden in the arena which is possibly why some horses fled during the unexpected noise test and the umbrella test. Unfortunately these influences were unavoidable as the data collection had to be fitted into these busy time schedules of the mounted units.

The fourth limitation is concerned with external influences on the days the tests were conducted, such as horses being exercised on the horse walker and the dogs from the dog unit barking. However these should have had little influence on the results as horses should be habituated to working in this environment as they face these influences on a daily basis.

The effect of past experience on responses

Interestingly Chamerlain (2004), Michaels (2004) and Warbey (2004), when studying breed differences in responses to these same tests, found better correlations to the farm the horse was raised on irrespective of breed. This suggests that the horses' prior experiences had more influence on their responses than any genetic similarities that may underpin behavioural responses. This may account for the findings here, where the animal's past experience of similar circumstances to the test conditions influenced their response more than any underlying personality, and so there was very little association between these responses and the horses' suitability for their roles.

This factor of past experience is a possible reason for the lack of correlations between the non-reactive standardised behaviour tests and suitability scores. On reflection, the unusual surface test was probably not a suitable test to use on these horses as it is likely that they may have been trained to cross such surfaces. This question was asked prior to commencing the tests, as it has been suggested by Christensen et al. (2008), that horses react differently to objects of differing colours, and it was ensured that a differently coloured plastic sheet was used to that, which previously the unit's horses may have been trained to cross. Despite using a differently coloured sheet, it was not possible to establish with certainty what colour plastic sheet the horses may have crossed in the past and if indeed the colour of the plastic sheet actually affects their behavioural response. The police horses were not overly reactive to the plastic sheet, they showed slight variation on the approach, most horses crossed after encouragement.

With the stranger approach test, the behavioural reactions of the horses could have been influenced by the horses' past experience. Most of the police horses stood still and looked at the human waiting for them to approach. This could be because these horses are taught to respect a person's body space and so waited to be approached. There were exceptions to this, which could have been due to horses having excess energy and taking advantage of their time loose in the arena. The lack of

relationship between this test and the suitability assessment is supported by Seaman et al. (2002) who found no association between the scores given by farm leaders on the horses' subjective behavioural score and its response to an unusual person. Hausberger and Muller (2002) also found the reactions of horses to vary between caretakers, so using different testers may have influenced the horses' responses to the tests.

This issue of past experience may explain why these tests are more suited for foals. The difference between the reactions to a novel object of untrained and trained horses has been demonstrated by Wolff et al. (1997) (untrained) and Le Scolan et al. (1997) (trained), who compared the results from their studies and found the untrained horses to be more reactive to the novel object. Le Scolan et al. (1997), states that this is likely to be because the horses in their study were used in games, where colourful objects were placed in the arena. This therefore led to a less reactive response, again demonstrating that these tests may be suited for untrained animals, which have not been influenced by past experiences and training. The police horses' responses to a novel object could have been influenced by past training, as although they are unlikely to have seen a lampshade, they do face many other novel objects when out on patrol and during training.

This is supported by Seaman et al. (2002), who suggest that in adult horses the use of standardised behaviour tests is limited, as the horses' background is often unknown. Horses frequently have a number of different owners and may therefore have been exposed to an array of various stimuli and

situations. The author continues to suggest that these tests would be of more use in foals, as demonstrated in past studies using foals and young horses. Visser et al. (2003), found a novel object test and unusual surface test to be valid in predicting show-jumping performance. They measured the foals' responses to a set of standardised behaviour tests and then assessed their performance as show-jumpers at the age of three. There were significant correlations between the horses' responses to a novel object/unusual surface and their show jumping performance. The horses used had been reared using similar methods, which meant their history, training and past experiences were alike and consistent and it may be that such tests are more suited in this context. These findings support the idea that standardised behaviour tests measure past experience, as tests used on these younger animals were more reliable and predictive of suitability of use.

The effect of learning and generalisation on responses

Horses have also been shown to have good spatial memory and can recognise when objects in their surroundings have been moved spatially (Hanggi, 2005). Research has been conducted to explore the extent to which horses' behaviour is sensitive to control by environmental cues, known as discrimination learning (Nicol, 2002). In these studies the horses learn that one stimulus, but not another will result in reinforcement (Hanggi, 2005), for example horses' have been found to discriminate between three push doors, with the correct one indicated as black and white (Mader & Price, 1982).

Discrimination learning can also be seen in feral horses who must learn to discriminate between group members and non group members (Heitor & Vicente, 2007) and they must also remember complex sequences to successive spatial cues to move between different locations (Nicol, 2002). Discrimination learning has also been used to study why horses startle at objects they have already encountered, it has been reported by Hangii (2005) that horses are able to transfer neural visual information, and therefore the reason the horses' startle to the same object is likely to be due to a difference of appearance when approaching, indicating a lack of generalisation.

A study by Christensen et al. (2008) also found a lack of generalisation in horses. In their study they habituated horses to six different objects of various shapes and colours. During the presentation of the objects they found no decrease in the horses' initial responses towards each successive novel object, indicating that habituation was stimulus specific.

This good spatial memory and poor ability to generalise means responses to standardised behaviour tests may not be suitable for measuring likely responses to non-test situations and this explains why there was little correlation between the horses' standardised behaviour test results and suitability scores. Any responses given by the horses to the test situation or in their everyday roles are likely to be stimulus specific and do not reveal any strong underlying predisposition to behave in a given manner. Potential exception to this is the reactivity to sudden noise/visual stimuli which may reflect the sensitivity of the horses' autonomic responses.

As prey animals horses are susceptible to stress, which is generally thought to inhibit learning ability therefore an aversive stimuli can produce a response which is not influenced by training (Mills & McDonnell, 2005) Police horses have a broad range of roles and are confronted with various sights and sounds on a daily basis; it is therefore likely that these tests cannot predict their suitability for police work.

Although these tests were not useful for adult horses, evidence suggests that these tests may be of use for foals, who are less likely to be influenced by past experience on responses. Therefore responses of naïve animals may better reflect personality and so suitability, which is supported by the research conducted by Visser (2002).

The possibility of reactivity as a predictor of suitability

Where moderate correlations were found between responses to standardised behaviour tests and suitability for use, these were all in reactivity tests.

This may be because these were all responses to unpredictable situations – sudden moving object or noise or person moving towards them and horses are likely to give automatic responses which are less influenced by past training and may better reflect the horse's underlying reactivity (personality). Such variation in reactivity may feed into the horse's suitability for some police horse roles. Although police horses have training to help desensitise them to various noises and moving stimuli, this training usually takes place

with a rider on board, who can control the horse's reactions and possibly give cues in anticipation of the stimulus. It is also noted by Christensen et al. (2008) that the riding situation differs substantially from a test situation, where the horse is trained to respond to the riders' aids. This is a possible reason why horses responded differently in the tests than they would when ridden, as reflected in the suitability scores.

The unexpected loud noise used in this study was unusual to the horses and the police horses had only been trained with open manual umbrellas, so the sudden opening of an automatic umbrella was also new to them. These environmental challenges were therefore novel to the horses and were less likely to be influenced by prior experience. This could be why more variation was shown by the horses during these tests. It is also difficult for horses to learn to expect these types of stimuli, even familiar ones, unless they are predictable. These reactivity tests probably do tap into the horses' basic sensitivity to sudden stimuli and not prior training.

A moderate correlation was found between a horse's postural expression and flight reaction to an unexpected noise and its suitability at community events, with less reactive horses being more suited. This finding relates to how a police horse should respond, as they need to be un-reactive around the public. An association was also found between a horse's tendency to return to touch a sudden moving object (automatic umbrella opening) and its suitability for officer training, with horses that touched the umbrella being more suited for this role. The touching of the umbrella indicates that these

horses are possibly more open to new experiences and are possibly bolder horses, making them more suited for officer training.

Moderate relationships were also found between a horses' flight reaction to an unexpected noise and its suitability for training officers, with more reactive horses being more suited for this role and the horses' flight reaction to a sudden moving object and its suitability for local patrols, public order and for training officers, with more reactive horses being suited for this role.

These results are unexpected as it would be expected that the horses more suited for these roles would be less reactive. A possible suggestion could be that normally the horses' behaviour is influenced by their rider and so their natural reaction was different to that normally observed by the officers. Another reason which will have influenced these findings is that only two of the 15 horses fled from the umbrella and these were two of the best police horses. Other factors are outlined in the limitations.

An association was also found between a horse's tendency to touch a sudden moving object and its suitability at community events with horses not touching the umbrella being more suited for this role. Again this finding is unexpected as it would be expected that the braver horses would touch the umbrella, again this is likely to be due to the two best police horses fleeing.

An association was also found between the horses' latency to touch a stranger and its suitability at community events, with horses' with a

decreased latency being more suited for this role. Police horses come face to face with a variety of people on patrol and therefore need to be confident around them. It would therefore be expected that horses which were more suited for community events would approach an unknown person quicker, which was demonstrated by these findings.

The findings from the current study are not supported by another study which assessed the use of reactivity tests in predicting temperament in therapeutic riding horses (Anderson et al., 1999). They found no relationship between the reactivity tests and the temperament scores of therapeutic riding horses, although they combined the reactivity scores, which made it impossible to assess each individual test. The horses were also restrained, which did not give the horses the opportunity to express natural behavioural tendencies.

Slabbert and Odendall (1999) have noted that startle tests should be handled with care and set up so the animals' are in an open space where they can escape and return to investigate the stimulus, these authors' found their startle tests to be successful in predicting police dog suitability.

If horses are given the opportunity to express a full range of behaviours, the variation of horses' responses in reactivity may underpin the variation in the horses' suitability and such tests may be appropriate for assessing suitability for roles in adult horses. This may be useful for therapeutic or riding school horses where a strong reaction to unexpected moving objects/noises would be a disadvantage.

7.1 Implications of the study

In practice the results from this study suggest that standardised behavioural tests are not suited to predict a horse's suitability for use as a police horse. Results from other studies such as Anderson et al. (1999), also suggest they are not useful for predicting therapeutic riding horses' suitability. The tests are on the whole not reliable or valid and are probably not worth pursuing in adult horses, whose history is unknown. Reactivity tests may be useful in situations where reactivity is a pre-cursor of training. This finding is disappointing, because valid standardised behaviour tests could have been used by the equine industry to aid in selecting horses suited for a particular job, which would have improved horse welfare and reduced costs and time in training horses unsuited for a particular job.

7.2 Future research

This study indicates that pursuing the use of behavioural tests as predictors of suitability for a particular job of adult horses is of limited value. It may be worth conducting more tests using different reactivity tests to see whether these produce the same good reliability and relation to suitability. As more valid findings have been established in studies assessing the use of these tests in young horses, future research should investigate whether the responses of foals can predict their suitability at a particular future job. Such work was undertaken by Visser (2002) on show-jumpers; future research could explore different roles with wider challenges. The experiences of these horses would need to be kept constant, to rule out individual differences in

behaviour due to training, bad experiences etc. This research would be time consuming as horses would need to be assessed as foals and then followed up a few years later. This may be useful for performance horses e.g. eventers, where foals are reared and brought up at the same establishment. Unfortunately, this would be difficult to achieve for police horses and riding school horses as they are often brought in as adults.