

## **Chapter 5**

### **Development of the 'suitability of use questionnaire'**

In this chapter a measure of police horse suitability was developed in order to assess the validity of the behaviour tests in predicting police horse suitability (Chapter 6). The literature review (Chapter 2) revealed that research aimed at assessing an animal's suitability in a specific role is limited, with the majority of studies reported within the canine literature. The most practical, valid and reliable method found was the development of a questionnaire, such as the one used by Serpell and Hsu (2001), who asked puppy-raisers to rate their dogs' typical response to common situations and events using a series of five point semantic differential type rating scales. They found their responses to be reliable and valid in assessing the suitability of a puppy as potential guide dog (See Section 2.2). As Serpell and Hsu (2002), had such success with their questionnaire, a similar method was developed here to allow horses' responses in situations relating to their work to be rated, giving a measure of how good the horse is at various aspects of its job.

#### **5.1 Questionnaire design**

The first task in designing the questionnaire was to find out what makes a good police horse by establishing how a horse should ideally react to common policing situations. This was achieved by interviewing an experienced mounted police officer and trainer from the Merseyside Mounted Police Section. This interview generated a list of roles and the officer

described how a horse would ideally behave in situations relating to those roles. The roles included:

- Local patrols – riding out from the stables: *situations* relating to this included riding through traffic, riding past street furniture both static and non-static, passing road-works, coping with sudden loud noises, riding out in close contact with other horses.
- Box patrols – travelling to area to be patrolled: specific *situations* relating to this included loading into and unloading from a horse box, and being in close proximity of other horses in the box.
- Night patrols – patrolling areas at night, e.g. city centre: specific *situation* relating to this included the ability to deal with working in the dark, such as dim street lighting and car head lights.
- Public order/riot training – used for events such as football matches: specific *situations* relating to this included large unpredictable crowds and sudden loud noises.
- Community events/school visits – public get to meet the horses: specific *situation* included lots of strange people petting the horses.
- Officer training – to become mounted police officers: specific *situation* included beginners learning to ride, so horses are required to be calm and well disciplined.
- Water – whilst out on patrol horses may be faced with some form of water: *situations* include crossing water, riding through/next to running water, and being splashed by a vehicle through the water.

From the above, a list of 19 questions was drawn up which enabled the horses to be rated in regards to situations relating to their job (see Appendix 7). The rationale for these questions can be found in Appendix 13. For the majority of the questions horses were rated on how forward straight and calm they remained in each situation (FSC).

### **5.1.1 Rating scales**

Ratings are commonly used for outcome (performance) criteria, e.g. employee performance, and they can reduce the subjectivity in observational and interview judgments by standardizing what is to be rated (Aiken, 1996). Five point rating scales of how FSC the horse remained in each situation, were used for the majority of the questions to assess horse performance. The number of points depends on the mental representation of the construct used (Albarracin, Johnson, & Zanna, 2005). When discussing the design of the questionnaire with an officer, it was felt that participants may have conflicting opinions on the horses' suitability, due to past experiences and therefore it was appropriate to use a five point scale, as a seven point scale may have given too much discrimination. These scales also offered a neutral point, as an odd number was used, so that the favorable categories were on one side and the unfavorable on the other (Oppenheim, 2006).

The scales allowed rating between a police horse's ideal behaviour and unacceptable behaviour. A mounted police officer advised that a police horse should remain FSC in all situations and this ideal became the positive end of

the scale for most of the questions. The opposite of this was an uncontrollable horse which became the other end of the scale for most questions. These were not applicable for all the questions, which are discussed in section 5.1.1.

## **5.2 Ethics (human)**

Before piloting the questionnaire it was necessary to obtain ethical approval for the use of human subjects in this study. A copy of the application and approval can be found in Appendix 8. Both Merseyside and Manchester police units completed informed consent giving permission for the employees at the units to participate in the study (see Appendix 9). The questionnaire used in this study was very low risk to the participants as it was completely anonymous and no personal questions were asked, as the questions were aimed at obtaining information about the horses' behaviour. Appropriate times to conduct the research were discussed with the yard manager, with the only inconvenience to the unit being time taken to complete the questionnaire. Both units were keen to assist with this research, as they thought it was interesting and it was thought that if the standardised behaviour tests proved to be reliable and valid, they could be a useful tool for the police units. The covering letter (see Appendix 10) and participant information sheet (see Appendix 11) outlined the research and it was made clear that by completing the questionnaire participants were giving their consent and they were free to withdraw at any time. The participation information sheet contains reference to a personality questionnaire, which was part of the wider research project linking personality and suitability.

### **5.3 Piloting**

The views were sought from yard managers and experienced officers from both units, who were asked to give their opinion on the questions, ratings and design used. It became apparent when piloting this questionnaire at Manchester Mounted Police Unit that this scale was not appropriate, as no police horse would ever be uncontrollable and the scale therefore did not reflect the range of potential ratings for the target horses. This led to the development of the new scale, which used percentages of time the horse remained FSC. The scale ran from the horse remaining FSC 100% of the time to the horse remaining FSC 60% of the time. This scale was advancing in 10% increments. A cut off point of 60% was used, as after discussions with the yard manager at Manchester Police Unit it was decided that it would be very unlikely for a police horse to be FSC less than 60% of the time. The positive end of the scale always remained on the left to avoid confusion by the respondent.

For most of the questions it was appropriate to use rating scales from FSC 100% - 60%, but for some questions these FSC scales were not appropriate. This included Q. 2.6, relating to the horse's ability to cope when in close proximity of other horses and Q. 2.10, concerned with the horse's reaction to other horses in the box. Instead of FSC, a percentage of tolerance was used for these questions, again ranging from the horse being tolerant towards other horses 100% of the time, to a limited tolerance of 60% of the time. Q. 2.7 asked how the horse reacts to being separated from other horses. Instead of FSC, the scale measured 'ease of separation', which ran from 'will

easily separate' from other horses 100% to 60 % of the time. Question 2.19 was concerned with the suitability of the horse as a schoolmaster. The five point rating scale remained in use for this question as it was concerned with the suitability of the horse for beginners. It was advised that some police horses were unsuitable and the scale therefore progressed from 'is suitable for beginners' to 'is unsuitable for beginners'. The final version of the questionnaire can be seen in Appendix 7.

#### **5.4 Target population/sampling strategy**

The target population for this study were the horses at the Merseyside and Manchester Mounted Police Sections. Both units completed questionnaires for all the horses in work at the time of the study, Merseyside completed N=26 and Manchester returned N=19, giving a total of N=45 responses. Sampling was not necessary as the questionnaires were completed for both units' entire population of horses in work at the time of the study.

#### **5.5 Procedure**

The population required to complete the questionnaires were two people who were familiar with the horse in question. These were officers or trainers who had worked with the horse and could give an accurate account of its behaviour. Questionnaires were supplied to the yard managers at the units, who distributed them to the appropriate members of staff. Participants were given a covering letter and a participant information sheet (see Appendices 10 & 11), outlining the details of the study and explaining the importance of not discussing their responses with each other. If people discussed the

questions, it may have influenced their answers, which in turn would query the reliability of the analysis. To give the units sufficient time to respond, the questionnaires were collected from the yard managers approximately two weeks later.

## **5.6 Analysis methods**

During data collection, two people rated each horse. This permitted reliability analysis of each of the 19 questions on the scale. The response ratings for each question were converted into scores (100% = 5, 90% =4, 80% =3, 70% = 2, 60% =1) and then inputted into SPSS for analysis. Weighted Kappa ( $k_w$ ) was used to assess the amount of agreement between two people. Altman (1991) provides a guideline to help interpret values, which can be seen in Table 4.4.1.

## **5.7 Results of reliability assessment**

Table 5.7.1 shows the results of the reliability testing of the items on the suitability questionnaire. These results indicate that all except the horse's response to being separated from other horses on patrol had at least moderate reliability (0.41-0.6), with 13 out of the 19 questions showing good reliability (0.6-0.8).

Table 5.7.1. Summary of the Weighted Kappa test (Kw) to assess the reliability of the suitability of use questionnaire. Bold values of Kw indicate moderate to good reliability between two independent raters.

<b>Item on questionnaire</b> <b>Horses reaction to:</b>	<b>Reliability</b>	<b>N</b>
Large crowds	<b>0.637</b>	43
Static street furniture	<b>0.617</b>	44
Non static street furniture	<b>0.617</b>	44
Heavy traffic	<b>0.747</b>	44
Road works	<b>0.636</b>	44
Other horses on patrol	<b>0.664</b>	44
Being separated from other horses on patrol	0.389	44
Being petted by strangers	<b>0.555</b>	44
Loading into a horse box or trailer	<b>0.661</b>	44
Unloading from a horse box or trailer	<b>0.693</b>	44
Other horses in the box or trailer	<b>0.708</b>	44
Oncoming traffic	<b>0.611</b>	44
Traffic approaching from behind	<b>0.544</b>	44
Patrolling at night	<b>0.655</b>	43
Loud noises	<b>0.566</b>	44
Crossing water	<b>0.667</b>	44
Being splashed by water	<b>0.510</b>	44
Running water	<b>0.532</b>	44
Horses' suitability as a school master	<b>0.71</b>	26

## 5.8 Development of suitability of use indices

Due to the large number of reliable items N=18, it was necessary to combine items into groups to reduce the risk of type 1 error, as multiple significance testing gives a high probability of finding a significant difference just by chance (Altman, 1991). Data were grouped by specific police horse roles, as discussed in Section 5.1. The groups were developed by adding up the situation specific scores for each of the police horse roles, a summary of which can be seen in Table 5.8.1. Once the scores had been added up they were standardised to a score out of 10 using the formula:

$$\frac{\text{Actual score}}{\text{Maximum score}} \times 10$$

A high score indicated a horse was more suited at a particular role.

Table 5.8.1. Summary of police horse roles and situations relating to them, the mean (SD) for the horses overall scores for each role and the number of horses used in the analysis are also presented.

<b>Police horse role</b>	<b>Situation specific questions</b>	<b>Mean (SD)</b>	<b>N</b>
Local patrol	Horses' response to heavy traffic, traffic behind, traffic in front, roadwork's, loud noises and other horses on patrol (tolerance and separate).	6.81 (1.79)	44
Box patrol	Horses' response to loading, unloading and other horses in the box.	8.93 (1.35)	44
Night patrol	Horses' reaction to patrolling at night	7.07 (1.87)	43
Public order	Horses' response to large crowds and loud noises	6.44 (2.19)	43
Community events	Horses' reaction to being petted by strangers	8.97 (1.89)	45
Training students	Horses' suitability as a school master	6.23 (3.41)	26
Water	Horses' response to being near to running water, crossing water and being splashed with water	6.25 (1.91)	44

The results from the questionnaire indicate that 18 of the 19 items generated reliable data, with only the horses' response to being separated from other horses on patrol generating unreliable agreement between two raters ( $Kw < 0.4$ ). This result could be due to many reasons, which may have included the officer having had a particular bad experience, in which a horse was unwilling to separate from another horse. Other reasons may include officers normally staying close to each other on their patrols and therefore not having greatly experienced this situation and so being unable to rate this item accurately, leading to differing opinions on how easily the horse separates from other horses. Two sets of questionnaires were completed for N=44 horses for the majority of questions. Only N= 26 were completed for the item 'horses suitability as a schoolmaster'. This is because Manchester Police Unit left this question blank, possibly because the people who completed the questionnaire were unable to accurately answer this question.

The scores from these items could be used to measure the suitability of individual police horses for the range of situations they are expected to cope with in their role as a police horse. The mean (SDs) are shown in Table 5.8.1, and their variation indicate that police horses perform better in certain situations than others. The standard deviations indicate that there are individual differences in a horse's reaction to the different situations and it is these variations in individual scores that were used in the next chapter to compare with the horses' behavioural responses to the standardised behaviour tests.