The role of observer individual differences in personality assessments of the domesticated horse: A novel application of Kelly's Repertory Grid Technique

DAŠA (DASHA) GRAJFONER

Thesis submitted for the degree of PhD in Psychology

The University of Edinburgh

2006

For my mum and dad, their unconditional love and support.
Draga mama in ata, ta naloga je posvečena vama: hvala za vajino ljubezen in podporo.

Declaration

1) This thesis was written by Daša (Dasha) Grajfoner.

2) Both studies presented in this thesis were all organised, performed, analysed and written by myself.

3) I hold the degrees of BA in Slovenian Language and Literature, BA in Social Sciences, MA by Research in Social Sciences (all from Ljubljana Unversity, Slovenija), and taught MSc in Applied Animal Behaviour and Welfare (Edinburgh University, Scotland).

4) Neither this thesis or the results from the two studies have been submited for any other degree or professional qualification.

Daša Grajfoner Date: 29.09.2006

THE UNIVERSITY OF EDINBURGH

ABSTRACT OF THESIS

Regulation 3.1.14

Name of Candidate: *Daša Grajfoner*

Address :

Postal Code:

Degree: *PhD*

Title of Thesis: **The role of observer individual differences in personality assessments of the domesticated horse: A novel application of Kelly's Repertory Grid Technique**

No. of words in the main text of Thesis: *60,533*

Consistent personality differences (e.g. shyness, sociability) among animals have been frequently reported by scientists. Such information used to be considered unscientific, even though it was widely used to describe animals and predict their behaviour (Hebb, 1946). Over the past three decades however animal personality has been studied systematically in various species (Gosling, 2001), using provided lists of descriptors, consistent with the five-factor model of personality (Costa & McCrea, 1992). Few researchers have allowed raters to produce their own list of descriptors (Dutton et al., 1997; Wemelsfelder et al., 2000), and none have investigated the potential influence of human individual differences on ratings.

The three main goals of this thesis were (i) to develop in-depth personality profiles of domesticated horses, (ii) to investigate the effect of personal backgrounds and attitudes of observers in their construction of these profiles, and (iii) to provide insight into the reliability and validity of the provided assessments of horse personality. The thesis starts with a literature review of personality psychology, psychometrics, animal and horse personality; followed by the introduction of the concept of animal-as-a-scientist. As methodology is an important part of the project, the repertory grid technique (RGT) originally developed by Kelly (1955), was employed as a novel method to assess horse personality. This method gives observers the freedom to generate their *own* descriptors, which allows them to integrate the totality of their experience with animals into personality constructs which are meaningful to them. RGT has previously been used for personality assessment in chimpanzees (Dutton et al., 1997) and for the assessment of individual styles of interaction in pigs (Grajfoner et al., 2002).

The experimental part of the thesis consists of two studies. In both, horse personality was assessed by groups of human participants, either familiar or unfamiliar with the horses. The familiar groups assessed the horses based on their past experiences. The unfamiliar groups watched short videos of horses interacting with a human. In the first study 44 female observers rated 21 horses from two stables. The results show a significant degree of agreement within the observer groups for 95% of horses. However, correlation of horse personality scores between the observer groups was not significant. The degree of agreement was higher when the observers were familiar not only with the horses but also with the descriptors (Adams-Weber, 1970). In the second study we further investigated the observed incongruity between familiar and unfamiliar observers. A novel object test was added to provide the unfamiliar observers with more information about the horses. Thirty four female observers rated 38 horses from three stables. Again, the degree of agreement was highly significant within the observer groups; however, correlation of the results between familiar and unfamiliar observers was only marginally improved. The degree of agreement was not consistently higher when the horses were rated on elicited constructs. Both studies showed that two personality dimensions, neuroticism and extroversion, are fairly robust in horses. The third dimension, agreeableness, was generated only by familiar observers. Contrary to expectations, observers' personality, empathy or emotional intelligence did not significantly affect the degree of agreement or how they rated the horses' personalities. Assessment of individual differences in horses using qualitative descriptors generated by the observers themselves is therefore not observers' self projection. These results make a significant contribution to the debate on anthropomorphism. The overall degrees of agreement between the observers indicates consistent reliability of the RGT throughout the observer groups in both studies.

Finally, the academic and practical implications of the study are discussed. On an academic level, individual differences in animals are of pivotal importance for understanding personality in the contexts of evolutionary, comparative and social psychology. On a practical level horses are, according to their individual differences, selectively used for different purposes: horse assisted therapy, racing and the police. Avenues for investigating the relationships between animal personality, performance and welfare should therefore be further explored.

Acknowledgments

As far back as I can remember I have been determined to find out more and to contribute to knowing more about what I believed in: that animals have emotions, personalities and individual integrity and that there are ways in which science can investigate this. First, I would especially like to thank both of my dream team supervisors, Dr. Elizabeth Austin and Dr. Francoise Wemelsfelder, for allowing me to expand my knowledge. Francoise has been not only my supervisor, but also an inspiring mentor, whose meticulous research and unique yet powerful theoretical perspective on animals have helped me to clarify my own ideas and contributed significantly to my research decisions. In the same way I would like to thank Elizabeth for her invaluable contributions to the methodology and experimental design of the project. I am particularly grateful for her constant support, enthusiasm and understanding throughout the time of my study. I feel privileged to have worked with two such distinguished scholars and good people. Secondly, I would like to thank all my participants, horses and humans – without them giving their precious time this project would not have been possible. Thanks goes also to Davy for his contributions to record the videos and to Pat for proofreading my thesis and for the words of encouragement. Thanks must also be given to the Faculty Group of Law and Social Sciences of Edinburgh University and the Overseas Research Students Awards Scheme. Finally, I would like to thank my parents who have always believed in me, and provided a never ending source of strength, especially during the last few months. I would not have been as productive without their love.

Contents

CHAPTER 1

An Introduction to the Project

1.1. Introduction

Do all animals behave in the same way and follow the same behavioural patterns or do they manifest differences in how they interact and relate to their social and physical environment? If there are differences in their behaviour are they consistent? Can we denote those differences as animal personality? What do we mean by animal personality? How do we define and assess it? Is the assessment reliable and valid? How many personality dimensions are there? Is it important for us to be familiar with the personality descriptors we use as well as with animals? Do our personality and ability to empathise affect the assessment?

Companion animals, principally dogs and cats, are animals we most commonly ascribe personality attributes to. They are expressive, usually retain adolescent behaviour and have plenty of interactions with us. There are also other species, e.g. horses, which we use not only for company: they also play an important role as sport, work and farm animals. Therefore the requirements for their behavioural traits are varied. The versatility of horses and lack of consistent horse personality research contributed to our decision to choose horses for this project.

The initial questions listed above were narrowed down to these:
1) How reliable is Repertory grid technique as the method for assessing horse personality? 2) What are horse personality dimensions? 3) Do observers' individual differences influence the assessment?

The main objective of this thesis is to answer these three questions by reporting and discussing the results from two studies. We chose Repertory Grid Technique (RGT) as the method for data collection. RGT gives observers the freedom to choose their own behavioural descriptors and rate the animals on them. RGT has previously been used to assess the personality of chimpanzees (Dutton et al., 1997) and pigs' styles of interaction (Grajfoner et al., 2002). However, it has not been applied to assess horse personality.

1.2. The importance of the research problem

The role of horses in human society is versatile. They are companions as well as sport, work and farm animals. As companion animals they have an emotional value and their role is mostly related to recreational riding and grooming.

As work animals, horses play a significant role in animal assisted therapy programmes, from riding for the disabled to equine assisted psychotherapy (Anderson et al., 1999). In a completely different field they are trained and employed by the police and military. In sport, horses are essential in polo, show jumping, racing and dressage. As farm animals they are still used for heavy labour and for meat.

In all these situations horses' suitability and performance are fundamental. Behavioural or personality traits - not just physical ones - will determine their success in a particular environment. Besides performance there are also welfare implications in placing the right horse in the right environment. Horses in therapy are required to be placid and calm, friendly and approachable, which may not be desired for horses in sport, where they are expected to be energetic and active. One of the outcomes of this project is to help create a uniform tool for reliable assessment of horse personality, a tool that could contribute to an improvement in horses' performance and welfare. Another important aspect of this project is its contribution to further understand the evolution of personality. Personality is a biological and social phenomenon. By discovering personality dimensions in animals, we can find more information about the biological and social basis of human personality and its development.

1.3. The novelty of the project

Finding a method that will allow observers to be involved in the assessment from the very beginning was essential. We wanted the observers to integrate the totality of their experiences with horses (Wemelsfelder et al., 2000). Repertory grid technique (RGT) as a method allows that. It enables the participants to be actively involved in the construction of the final scoring list, which may increase the reliability of the assessment (Fransella et al., 2004) and permits the use of more advanced statistical tests. The technique allows the collection of both qualitative and quantitative data.

During the first stage of our studies we created a pool of horse personality adjectives, generated by experienced and inexperienced observers. This information can in the future help to create a horse personality questionnaire and test its reliability and validity. The validity of horse personality dimensions can also be explored by cross-validation with established human personality tests, e.g. the NEO five factor personality questionnaire (Costa and McCrae, 1992).

A further novelty of our project was assessing the personality, empathy and emotional intelligence of observers as well horses, and exploring how observers' individual differences influence their assessment. In theoretical debates, assessing animal emotions and personality is often dismissed as anthropomorphic (Crist, 2000). By exploring the relationship between observers' personality, empathy (Paul, 2000) or emotional intelligence (Schutte et al, 1998) and their assessments of horses, we will be able to contribute to the anthropomorphism debate with some empirical data.

The final innovation of the thesis is the theoretical idea of the animal-as-scientist. It is presented as a perspective on animal behaviour that integrates intra- and inter-species interactions. The model provides a ground for mutual animal-animal, human-animal and human-human assessment of behaviour, making predictions and anticipating future events and behaviours for both animals and humans based on their previous experience.

1.4. Summary of the structure of the thesis

The thesis provides an introduction to animal personality: by exploring psychometric aspects of the relevant personality theories and by reviewing animal personality studies. This is followed by the presentation and discussion of the results from our two studies.

In Chapter 2 we aim to identify the requirements for a comprehensive personality theory. We also look at the main psychometric issues, validity and reliability, present in human and animal personality research (Gosling, 2001).

In Chapter 3 we highlight trait and phenomenological personality theories and briefly discuss biological models. The latter highlight the evolutionary perspective, important for its contributions to understanding the development of human personality and for considering personality as a continuum across species rather than there being a qualitative cut off point between humans and other animals. The second group are trait related theories. In the study, participants have to generate words that best describe horses' behaviour. Human as well as animal trait theories address the question of how many personality traits and dimensions there are. Is it enough to have three, five or even 16 dimensions (Costa and McCrae, 1995)? Most animal personality studies, including ours, make reference to three or five factor models of personality (Gosling, 2001). Trait theories are also relevant as we measure observers' personality using the NEO personality questionnaire (Costa and McCrae, 1992). Phenomenological personality theories on the other hand focus on a person's perspective and their phenomenal field (Kelly, 1955; Rogers, 1989). Interest in the construction of reality by an individual is personal or idiographic and provides mostly qualitative data. The basic premise of phenomenological theories is subjects', including human and non-human animals' active or intentional involvement in their interactions with their social and physical environment (Rogers, 1989; Wemelsfelder, 1999). They construe their experience of surroundings and create a perceptual or phenomenal field that is crucial for the development of their personality (Rogers, 1989). The phenomenological perspective will be discussed throughout the thesis.

In Chapter 4 we review animal personality research from a methodological perspective, as the novel application of RGT is an important part of our project. We also look at how many and which personality dimensions are identified. In the first part of the chapter we review personality research on primates and on farm and companion animals; in the second we focus on horse personality and temperament studies and look at the relationship between personality and performance in horses.

In Chapter 5 we discuss the idea of subjectivity and inter-subjectivity. Kelly's theoretical concept of a person as a scientist is embedded in his Personal Construct Theory (Kelly, 1955) and represents a fundamental starting point for the introduction and development of our concept of animal-as-scientist. Personal construct theory states that in the same way as scientists, we all make hypotheses and test them out. We anticipate future events by construing their replications. The process is not limited to cognitive information processing, but is more holistic and includes emotions. We use those anticipations and constructions to predict and replicate future events and interactions with others. Interactions are not limited only to humans; they are a mutual act, shared with other species (Sanders, 1992). Animals, whether they are involved in an inter- or intra-species interaction, have to be able to predict and anticipate the behaviour of others to have interactions that are meaningful and/or desirable (Boeree, 2006). Our theoretical chapter introduces two ideas: the concept of animal-as-scientist and the notion of human-animal interaction as a mutually constructed, initiated and shared act.

Chapter 6 introduces Repertory grid technique as a method for assessing horse personality. Kelly used the method mostly in a clinical setting; however, it has been applied to other areas like food or drink assessment (see Fransella et al., 2004). The method has three main stages: elicitation, interviews and scoring. During elicitation the participants generate bipolar constructs like "shy-bold" and during the interview they elaborate, rephrase or explain their constructs. In the last stage the observers rate a particular object or subject on the constructs they have provided. Repertory grid technique fits well with the idea of animal-as-scientist which we discuss in Chapter 5.

Chapter 7 reports the methods and results from the first study. This is the first attempt to use RGT for assessing horse personality. We have two groups of horses and two groups of observers, familiar and unfamiliar with the horses. The unfamiliar observers are further divided into participants with or without previous experience with horses. They both assess the horses by watching short videos of horses interacting with a human. Familiar observers only rate their own horses on the basis of their previous experience with them. The observer groups are treated as separate groups. As a result we end up with four different scoring lists, one per group. The results from the first study show that familiar observers have an overall higher degree of agreement. There are also no significant differences in the degree of agreement between the two unfamiliar groups. All observer groups generate two personality dimensions, labelled as neuroticism and extroversion. The familiar observers generate an additional dimension, which is not very well formed, but labelled "agreeableness". We find no consistent correlations of horse personality scores between familiar and unfamiliar observers. The reason for this might have been their different methods of observation or the fact that the unfamiliar observers did not have enough information about the horses.

In Chapter 8 we report the results from the second study. The experimental design is modified. First, we increase the number of horses; secondly, the unfamiliar observers are not split into experienced and inexperienced groups; thirdly, a novel object test is added to the video observation, so the unfamiliar observers have more diverse information about the horses; fourthly, the familiar observers are asked to rate each horse on their overall performance, using a single item scale. And finally, the scoring lists are further unified: we compose only two, one for familiar and one for unfamiliar observers. The results showed that familiar observers have a higher overall degree of agreement. In comparison to the first study, familiar and unfamiliar observers show a tendency to agree with horses' personality scores. We discover three personality dimensions, labelled "neuroticism", "extroversion" and "agreeableness". We also explore the relationship between the performance and personality profiles of the individual horses.

Chapter 9 reports the observers' personality, empathy and emotional intelligence results from both studies. The observers fill in the NEO personality questionnaire (Costa and McCrae, 1992). In the first study they complete the Animal Empathy scale (Paul, 2000) and in the second the Emotional Intelligence scale (Schutte et al, 1998). There are no consistent differences in the degree of agreement between more and less empathic observers, nor are there any differences between how they position the horses along the two dimensions. In the second study, observers high on the emotional intelligence scale tend to have a slightly higher degree of agreement. Neither observers' emotional intelligence nor their personality scores, however, have any significant impact on their positioning of the horses along the two dimensions. We were expecting that observers with higher scores on the neuroticism or extroversion dimension might have a different perception of horses' personality than the observers who score low on these dimensions. However, the results do not confirm our prediction.

In Chapter 10 we discuss the results. Neuroticism and extroversion are personality dimensions present across different species, including horses. Personality scores produced by familiar and unfamiliar observers are not significantly correlated. That may happen either because there are two different ways of data collection, using video or previous experience, or because of the different levels of familiarity with the horses. The degree of agreement is significantly higher within the familiar group of observers. Familiarity with the words however, is not as influential for the degree of agreement as we expected. We also briefly discuss the relationship between horses' personality and performance, and we recommend further investigation. Observers' individual differences did not have a consistent and significant impact on their assessment of horses. It may be that the observers do not project their own personality on horses, but are just rating their observable behaviour. The results suggest that more empathic people and those with higher emotional intelligence scores do not perceive horses differently. They may be more observant, but they still see the same behaviour and rate it consistently. We discuss the implications of those results. Assessing the qualitative aspects of animal behaviour is readily seen as subjective and anthropomorphic in the animal behaviour scientific community. Our study addresses this issue empirically. The results do not show a consistent occurrence of observers' personality, empathy or emotional intelligence being projected onto their rating of the horses.

1.5. Conclusions

In this project we formulate, address and answer the three questions stated at the beginning of this chapter. We identify personality dimensions in horses; test the reliability of the repertory grid technique, and explore how observers' personality, empathy and emotional intelligence influence their assessment. We introduce the theoretical idea of the animal-as-scientist, and make recommendations for further research.

CHAPTER 2

An Introduction to Personality Psychology

2.1. Introduction

Any attempt to address personality assessment, either in humans or animals, should include two components: theory and data collection (Pervin & John, 1997).

The theoretical part elaborates questions like what personality is, how it is structured and which factors affect the structure. Building up a structure, or developing a personality, applies to humans as well as animals. Genetics and social environment both influence our behaviour. Some of this behaviour is consistent over time and situations. The notion of consistency in our behaviour is an integrative part of most personality definitions. Pervin and John (1997) have proposed a broad, general definition, which is acceptable for most personality psychologists: "Personality represents those characteristics of the person that account for consistent patterns of feeling, thinking, and behaving." (ibid, p. 4).

The definition of personality can of course be broad as the one above. However, most theories have invented their own explanations that are mostly based on their philosophical frameworks. Even though it is widely accepted that personality is determined by both biological and social factors, different theories emphasise one aspect or the other. Social theories try to explain individual differences by external factors (Mischel, 1976; Hampson, 1995), while biological theories try to do the same by accentuating genetic and physiological components. However, no matter which outlook the theorists are coming from, for a personality theory to be comprehensive, they have to address various questions: about the nature of behavioural consistency over time and situation, the relationship between emotions, cognition, awareness and behaviour and even the connection between the concept of self and behaviour (Pervin & John, 1997).

To test our theoretical models and collect the data, we need to decide not only what our operational definition of personality is, but also what are the instruments of assessment or methods of data collection. Reliability and validity are the main psychometric issues that will be addressed in this chapter. We will look at different methods of data collection, particularly case studies and questionnaires or rating scales.

All those issues, theoretical and experimental, are relevant for animal personality research (Gosling, 2001). The goals of this chapter therefore are firstly, to address the most important theoretical and practical issues in personality theory, secondly, to review and compare the levels of explanation and lastly, to address the reliability and validity of the assessment. The major paradigms in personality psychology will be addressed and a transition to the relevant theoretical approaches discussed in the next chapter will be made. We will make reference to our framework and place it in a wider contexts.

2.2. The theoretical issues of personality psychology

Different approaches in human psychology have influenced the development of personality theories. Phenomenologists, psychoanalytical theorists, trait, biological and social psychologists have all attempted to define what constitutes a person or personality. Phenomenologists do not speak about personality; they focus on the development of self and self-actualization. Variation between the personality models or levels of explanation exists however not only between, but also within each theory. For example there are differences between phenomenological (Rogers, 1989; Kelly, 1955) and trait personality models, and there is a significant variation within each group of models as well (Eysenck, 1952; McCrae & Costa, 1987). Beside those different theoretical models of personality which will be addressed in the next chapter there are general issues any personality theory, human or animal, has to address (Gosling, 2001).

Animals like humans, show differences in the way they behave (Cooper, 1998; Gosling and John 1999). Some behaviours are consistent and allow us to make assumptions, observations, explanations and predictions of their future behaviour and maintain successful interactions (Hampson, 1995). Personalities differ not only in quantity but also in kind (Birch and Hayward, 1994), e.g. in addition to different degrees of calmness, we also have different dimensions: agreeableness, extroversion, neuroticism etc. Scientific approaches to exploring those differences are based on various theoretical perspectives, which involve efforts to systematically explain the differences between individuals and their everyday interactions (Kerlinger and Lee, 2000). Personality psychology therefore emphasises various psychological processes, including cognition, moods and emotions as an integrative description of a person as a whole. The purpose is to create a coherent picture of the individual's way of functioning, interacting and behaving. Personality research could be defined as an area of research that is concerned with how individuals as a whole differ from each other or indeed are similar. It depends whether the theory is idiographic (focusing on differences) or nomothetic (focusing on similarities and grouping individuals).

2.2.1. What makes a good personality theory?

For some psychologists our everyday assessment of others and a scientific approach to personality research are not fundamentally different (Kelly, 1955, see chapters 3, 5 and 6). For others this is not the case and a scientific study of personality has a different set of parameters and requirements (Eysenck, 1952). In general, the task of a good personality theory is to build up, test and validate the model it has created. The models need to be explicit and terminology consistent. Data collection should have a clear experimental design that is valid and reliable, has the ability to systematically test predictions or hypotheses and fits into the theoretical model (Rosenthal & Rosnow, 1991). In addition, a personality theory also needs to define its aims and constraints and provide operational definitions, essential for developing experimental methods (McCrae & Costa, 1997, 2000).

In human as well as animal personality theories there are basic questions that can be condensed into three major groups (Pervin & John, 1997): What are the characteristics of an individual and how they are organized; are those characteristics genetically or environmentally conditioned, and how is their interaction manifested in the individual; and finally why do the individuals behave in the way they do?

Animal personality studies mostly do not consider theoretical issues in that depth and focus on the experimental part instead. Chapter 4 demonstrates that animal personality usually takes over the basic principles of trait or biological models in human personality. However, if we attempt to approach animal personality consistently those questions should be addressed. They can be translated into four areas which a personality theory should cover (Pervin & John, 1997):

1) Structure refers to stable aspects of personality. Different theories have different elements that constitute that stable structure. They can be traits, types, response, expressions, constructs. Those elements are organized usually in a hierarchical way. In trait theories traits are represented by many qualitative descriptors e.g. "bold", "serious", "interested" etc, and can be clustered into types, factors or dimensions,

e.g. neuroticism, extroversion. Most animal personality studies do not cluster traits, but if they do, they are most likely comparable with human trait factors (Gosling, 2001). In comparison to traits which are possessed to a certain degree, type is a categorical concept (Birch and Hayward, 1994).

2) Process refers to motivational concepts. They are relevant to animal personality as animal behaviour science includes study of motivation as one of the main components. Pervin and John (1997) list three major categories: pleasure or hedonistic motivational concepts, self-actualisation and cognitive motives. Pleasure seeking and avoidance of pain is frequently studied by animal behaviour scientists, while self actualization has been included in discussion only by Wemelsfelder (1985a, 1985b, 1993, 1997a, 2001). On the other hand neuro-psychology and evolution of emotions has been studied with wider implications for animals (Darwin, 1872; Panksepp & Panksepp, 2000). The concept of cognitive motivation emphasises the ability to understand and predict the behaviour of others. Theory of mind is represented in animal behaviour studies and has been explored mainly by primatologists, but also other ethologists (see Griffin, 1992).

3) Growth and development clearly determine personality. They can be genetic and environmental. Even though the nature-nurture debate is still active, it has been generally accepted that both nature and nurture equally contribute to the development of personality.

4) Behavioural change and psychopathology are suggested to be a part of a complete personality theory (Pervin & John, 1997). It may be unusual to apply this terminology to animals; nevertheless it is quite common that some animals cope better with stress and interactions than others. The study of animal welfare should be particularly concerned with this aspect of individual differences. It is clear that the welfare of not all animals will be compromised and some animals may suffer more than others. Pet behaviour counselling is mostly concerned with this issue and psychopathology is addressed using mostly "cognitive behavioural therapy" (Overall, 1994, 1997).

2.2.2. Main issues in personality theory

Even though a personality theory will cover these four areas, there are questions constantly emerging and a theory can be assessed in accordance to how well it addresses and answers them (Pervin & John, 1997).

Human personality psychology will normally address the question of what it is that constitutes being human. The theoretical framework established in this way will influence personality theories. The same question can be asked about animals and the philosophical perspective will determine animal personality theories in the same way as they determine human personality theories. Therefore it is important to propose a philosophical view on which our animal personality theory is based (Wemelsfelder, 1993, 1997a, 1999a). The philosophical view of the person underlines the operational definition of personality and consequently the choice of what and how is measured. The basic distinction is between the reactional and the interactional. The first presents the person as an organism responding to external stimuli, driven by internal forces, while the second group are theories that define the person as an entity that interacts with the environment and with others and forms their own perspective and personality through those interactions.

We have already mentioned the discussion about internal and external factors of behaviour. There are three main perspectives: situationalism and behaviourism are more extreme, while the third concept, intearactionalism, appears as a response to the two. The person-situation discussion seems to echo the wider nature-nurture debate. Although there is a consistent dispute between trait theorists (Matthews and Deary, 1998) and situationists (Mischel, 1976) the extreme positions are rare. The most harmful criticism to trait theory for example would be to challenge their claims about behavioural consistency. Mischel (1976) argues that situations modify behaviour and that total behavioural consistency may be an indicator of poor adaptation. A situation might cause an individual to behave in a certain way and consistency comes not from within the person but from the situations. Mischel's arguments have led to more extreme positions defining behaviour as largely determined by situation. However, both Mischel and trait theorists do acknowledge that a combination of both is important, and that some traits are more consistent than others. Interactionism is a more elaborate fusion of both positions.

An additional view on consistency of behaviour is offered by phenomenological approaches (Rogers, 1989; Birch & Hayward, 1994). They focus on a person's own perception of consistency of some of their traits. A study by Bem and Allen (1974) suggests that the participants who perceived themselves as consistently friendly were rated like that also by their peers as opposed to participants who believed they were inconsistently friendly. Self assessed consistency of behaviour has also been widely acknowledged.

A theory can be idiographic or nomothetic. Idiographic theories emphasise the uniqueness of individuals and differences between them. The main sources of data are case studies. Nomothetic theories on the other hand tend to group individuals, emphasising similarities. The first group is represented mostly by phenomenological theories (Kelly, 1955; Rogers, 1989; Allport in Goldberg 1993). Trait theories are best described as nomothetic (Eysenck, 1952). An idiographic theory will normally build up a detailed profile of one individual at a time.

The concept of self is integrated in an individual's behavioural patterns. Behaviour expresses the unity, pattern and organization of a subject. Personality is therefore determined by how we feel and by self-concept (Pervin & John, 1997). Self-awareness is an important part of our phenomenological or subjective experience. Our behaviour is considerably influenced by how we feel about ourselves. As a consequence the concept of self is often an expression of integrative aspects of our personality. Although the importance of self in personality seems to be undisputable, theories differ in the degree to which they address the issue. Most emphasis on the unity of behaviour is in the idiographic personality theories of Rogers (1989) and Kelly (1955). The first argues that the person strives for self actualisation; for the second the constructs associated with self play an important role in a person's functioning. The concept of self is much less pronounced in nomothetic trait theories

(Funder, 1997). While some trait theorists admit that organization of traits into a coherent sense of self is an under-researched area (Matthews and Deary, 1998), others include self-concept in their models of personality (McCrae and Costa, 1997), although do not discuss how traits could relate to qualitative personal experience. The reason is of course in the underlying philosophical differences. While idiographic theories are based on many observations of an individual, nomothetic theories are concerned with exploring particular variables, e.g. a trait or a dimension, in a group of individuals. An idiographic or holistic theory is interested in exploring the relationship between the behaviour, feelings and thoughts of an individual. A nomothetic or reductionist theory groups individuals; it reduces them to a single point in multi-dimensional space. The ultimate goal is to study phenomena systematically and build up from simple to complex. The self or organization of behaviour becomes important when there is enough knowledge about the parts.

Consciousness or self-awareness is yet another topic on which personality theories differ. Trait theories use self-report questionnaires and peer or target ratings (Costa and McCrae, 1997) as instruments of assessment. Reliable self-assessment implies that we are self-aware. In a similar way phenomenological methods are based on individual subjective experiences. One of the problems facing personality theorists is how to conceptualise the role of varying states of consciousness in the individual's personality. We are not always aware of factors that influence our behaviour. How do we conceptualise that fact? Phenomenologist Rogers (1989) connects the unawareness to his idea of self-concept and possible denial of experiences connected to it. Kelly (1955) on the other hand suggests that one or both poles of a personal construct may be submerged and are unavailable to awareness. Trait theories in general do not address the concepts of consciousness or awareness.

Past, present and future behaviour all determine personality. That the present is an important determinant of behaviour is not disputable in any of the major personality theories. However the theorists are not so congruent as to whether the influence of past experiences or anticipation of the future have more influence on present behaviour and personality. For existential psychologists, present experiences and factors operating now have the most relevant influence on behaviour. Behaviourist and psychoanalytic theorists emphasise past behaviours and experiences as most significant in reshaping the present actions of an individual. Kelly accentuates the importance of explaining the meaning of past and present experiences by their construing and anticipating future events. Understanding the inter-relation between past, present and future is an important task for all personality theories.

2.3. Experimental Issues Concerning Personality

At the beginning of this chapter we mentioned that that even though we are all personality theorists and underlined procedures of assessment are similar, personality psychologists are more rigorous regarding consistent use of terminology and systematization of concepts. When it comes to data collection the situation is very much the same. As Kelly (1955) suggested with his Personal Construct Psychology, we all go through the same process of testing and confirming or rejecting our expectation or anticipation of how others are and how they behave (Kelly, 1955). Scientific data collection goes through the same stages, but is more open to systematic examination (Pervin & John, 1997). Collecting information is more rigorous and a great deal of attention is paid to consistent description of methodology so that the design can be replicated by others. In this section we will describe and discuss two groups of issues relevant to animal personality theories: first, we will identify different categories of data we can collect and then we will address the issue of reliability and validity (Gosling, 2001).

2.3.1. Sources of data

Four different types of information or data are used to assess personality in humans. The categorization was first formulated by Cattell (1957) and later on adapted by various theorists with abbreviations as LOTS (e.g. Pervin & John, 1997) or LIST

(Funder, 1997). Application to animals will require modifications. Animal behaviour science has developed its own categories of data collection (e.g. Martin & Bateson, 1998), although assessing animal individual differences does require further examination of data sources as categorised by human personality researchers. We will explain the two acronyms and relate them to animal personality.

The L stands for real life data on how the person is faring in life, e.g. number of injuries, education and their performance. This source of data will be relevant for assessing animal welfare (Webster, 1994).

The I or O stand for informant (Funder, 1997) or observer (Pervin & John, 1997). It refers to the data which is gathered in the form of judgements by knowledgeable individuals. The assessment of personality from O data is based on observations in real life and provides in-depth information. The judges or assessors have to integrate the variety of information they have about the individuals being rated. This source of data has been largely used for the assessment of animal personality (see Gosling, 2001) and has potential in animal welfare science (Wemelsfelder et al., 2000). The advantages of O data are in both the quantity and quality of information we can collect from a familiar rater. Therefore O data is not just a collection of different immediate or prior behaviours the rater has seen. The information is situated in the context and integrated in the observer's construction system (Kelly, 1955). According to Funder (1997) there are a few disadvantages with the observer data. In human personality assessment there is a limited amount of information a rater has. We tend to live in separate compartments and put forward different aspects of our personality in different contexts. At work we behave differently than at home or when we do sports. In animal personality research we may avoid this issue, as animals usually do not adopt different roles that would be drastically separated. The second disadvantage is that the raters are usually biased, which means that they might not like the individual they are assessing. This can be a problem in animal personality assessment as well. To minimize the error we should design experiments with a group of raters and calculate inter-rater agreement. Finally, Funder mentions errors in judgement, which can be an issue we need to address in animal personality research as well. Some individuals and some behaviours are more expressive and easier to assess than others. Here as well we can use inter-rater reliability to avoid this problem, although the raters could be collectively wrong.

S or self data are the simplest and easiest way to collect personality data from individuals who can make use of self administered questionnaires. This self assessment has an advantage as we know ourselves the best. We have access to our own expectations and how those expectations influence our behaviour. The disadvantages are obvious. Some individuals may not provide information about themselves or change it according to predicted expectations. They may want to present their personality in the most desirable way and minimize the traits which are least sought-after. Other individuals, including small children and animals are usually not able to give relevant information about their personality. They either lack insight or do not have the verbal or mental capacities to comply with the test instructions. In that case a different type of test should be devised to adjust the method to the abilities of the subjects in question (Funder, 1997). In animal welfare Dawkins (1980, 1988, 1990, 1993) developed an instrument of assessing animal preferences, which is a form of self-report data. According to Dawkins animals are motivated to do something and the experimental designs focus on how hard and how long animal will persist to achieve the goal.

The last group of data are T or test data. In comparison to I or O data this group of information is supposed to be non-judgemental and quantitative (Pervin & John, 1997). It is based on direct observations either in the real life context or in designed experimental settings where the subject is put in a highly controlled situation. The advantage is obvious – the manipulation of the context is higher and therefore the behaviours we want to observe can be triggered more easily and quickly. The information gathered is not filtered with judgements and therefore satisfies the experimental parameters of the quantitative, reductionist scientific context. One of the major problems of test data is the interpretation of results, which are normally linked to the same theoretical framework nurtured by research with similar experimental designs.

All types of data have their own positive and negative aspects. The choice depends on the research question, the theoretical background and the way we want to validate the results of our research.

2.3.1. Reliability and validity

Reliability and validity of the data collected are the two key psychometric concepts that are inevitably integral to all personality theories and experimental designs. While reliability can be simply denoted with the question "Can we get the same results more than once?" and refers to the consistency of a measuring instrument, validity refers to the meaning of that measurement and asks "Does the data mean what it says?" In general a valid measure must be reliable, but not vice versa. It is possible to have highly reliable data that is not valid.

Synonyms for reliability are dependability, stability, consistency, predictability. All psychological and behavioural measurements are more or less variable from one to another occasion. The smaller the error of measurement of the measurement the more reliable it is. Kerlinger and Lee (2000) suggest two approaches to the definition of reliability:

1) If we measure a personality dimension with the same or comparable instruments a number of times we will have the same or similar results. This implies definition in terms of stability, dependability and predictability. In trait personality research, stability is sometimes separated from reliability and defined as test-retest correlation of the scale over a period of time (Matthews and Deary, 1998). The distinction is supported by the argument that personality will change slowly over time, which is an unavoidable process and does not imply low reliability.
2) Another way of defining reliability is to look at how much error of measurement there is in a measuring instrument. The error can be systematic (scores obtained can be all higher or lower due to an unreliable scoring scale) or random (scores will not

have any definite pattern of error). Therefore reliability can be defined as the relative absence of errors of measurement in a measuring instrument. Experimental designs for assessing animal personality usually include a group of judges rating the animals. Inter-rater agreement signifies internal reliability. Internal reliability is proportionate with the degree of agreement: the higher the degree of agreement, the higher internal reliability.

Validity on the other hand refers to the accuracy of the measurement. Are we actually measuring what we think we are? With measuring a physical property the congruence between a measured variable and the measuring instrument is expected to be close. However, when measuring personality we usually have to invent indirect instruments of measurement. The definition of this again depends on the theoretical context of our experimental design. Kerlinger and Lee (2000) describe several forms of validity:

1) Content validity refers to how representative the topic or content of a measuring instrument is. For example does the scoring scale represent personality traits or emotions? The dilemma can be addressed by looking at the theoretical context on the basis of which the instrument has been developed.
2) Criterion related validity involves comparing our personality scoring scale with one or more external criteria which are known to measure the attribute under study. There are two types of criterion related validity, concurrent and predictive. The latter for example involves the future performance of the criterion.
3) Construct validity is the third type of validity (Cronbach and Meehl, 1955) and has a prominent philosophical basis. It links psychometric and theoretical notions by gathering together as many different measurements of a construct as possible (Funder, 1997). It answers the question of which factors or constructs are behind test performance (behaviour, style of interaction, dominance) and refers to the properties being measured. The construct validity depends on relevant empirical and theoretical knowledge (Cooper, 1998; Matthews and Deary, 1998)

Reliability and validity are often confused. Even though the concepts differ, some theorists indicate that the distinction between the two is not clear and that they represent only two aspects of one phenomenon – generalizability (Cronbach et al., 1972). As mentioned above, while reliability refers to accuracy of the measuring instrument validity refers to accuracy of the measuring of that specific property (e.g. personality or intelligence). If the reliability is high we do not have any information about what we are measuring – we only know that we are measuring something consistently. Reliability is lack of distortion and therefore represents the precision of a measuring instrument. Animal personality data usually indicate high internal reliability between the judges (Gosling, 2001). The validity however is not often addressed. In our experimental work the reliability of horse personality assessment was investigated by correlating horse personality scores produced by familiar and unfamiliar raters. Familiar observers were expected to achieve a higher degree of agreement due to their intimate knowledge of individual horses. We also expected familiar and unfamiliar observers to agree on their ratings. The latter would confirm reliability and validity of the assessment. The results will be presented and discussed in the following chapters.

2.4. Conclusions

In this chapter we introduced general psychometric issues that are an integrative part of any model of personality. Theoretical and experimental issues are often not addressed in animal personality research. Discussing them provides a structure for future experimental designs and systematises what has been done so far (Gosling, 2001). The issues dealt with in this chapter represent an introduction to the review of the most relevant personality theories in Chapter 3, which are followed by the review of animal individual differences in Chapter 4.

CHAPTER 3

Trait and Phenomenological Theories of Personality

3.1. Introduction

In the previous chapter we explored psychometric concepts relevant for human and animal personality theories and research. In this chapter we will look at two groups of personality theories, trait and phenomenological, and their definitions of personality. We will also briefly discuss the biological theory of personality. The first reason for doing this is to trace the development of the idea of main personality concepts and dimensions discovered in animals. The second reason is to emphasise the integration of both, as has been done in our project. Trait theories offer strong scientific, biological and nomothetic perspectives on individuals, while phenomenological theories are idiographic and emphasise individuals' perspectives and constructions, including the idea of self development and actualization. We will try to show that even though the two concepts may not be perceived as compatible (Matthews and Deary, 1998), they certainly can be integrated to provide more information about individuals. Each approach has several models of personality. We will look at them from an integrative perspective, bearing in mind animal personality. We will also discuss the limitations of each approach and their personality models. The purpose of this chapter therefore is to highlight the most relevant theories and their main problems from an animal personality perspective.

We will first define each theory and explain its relevance for the project. We will also highlight their definitions of personality and discuss the differences between personality, mood, states, activities, emotions and temperament. In human personality research the psychologists have tried to make clear distinctions between them (Eysenck, 1952; Allport in Goldberg, 1990), while in animal personality research the terms can be used inconsistently. Finally we will address the validation issue of target rating relevant for our research.

Personality theories make their contribution to understanding individual differences by looking at the formation of our knowledge about individuals and how they differ from or are similar to each other. The two groups of theories we have decided to look at are perceived to be quite different: trait and biological theories are defined as nomothetic which means that they focus on grouping individuals according to a limited number of personality traits and dimensions. Their personality models are usually descriptive and they aim to describe those dimensions from a social and biological perspective (Eysenck, 1952). Phenomenological theories on the other hand are idiographic and focus on how individuals differ from each other and what makes them unique. Another relevant group of theories are social constructivist (Hampson, 1995), which transpose the formation of personality from within a person to the interactive space between subjects (Crossley, 1996). Even though the models comply with different theoretical frameworks they are not mutually exclusive. There has already been a degree of integration at an early stage of trait theory development by Allport (Pervin and John, 1997), who was a trait theorist but emphasised individual traits and their importance for personality. Both groups are relevant to our project. Animal personality research usually adopts a nomothetic trait context. However there has been a development in an idiographic direction as well, especially in case studies based on human-animal interaction (Sanders, 1992, 1993).

In contrast with traditional behaviouristic approach to animal behaviour a trait oriented approach will assess qualitative aspect of behaviour, and allow application of the method that will give the observer the freedom to generate their own constructs (Dutton et al., 1997; Wemelsfelder et al., 2000). However, the goal is still to produce dimensions and group animals along those dimensions (Wemelsfelder et al, 2000; Grajfoner et al., 2002). This research however can also provide profiles of individual animals on the basis of two or more dimensions. Idiographic designs are mostly limited to case studies. For humans an ideographic approach is mostly applied in clinical settings. In animal behaviour, there have been only a few attempts to ideographically assess personality in pythons (Dutton, 2004) or dogs (Sanders 1992). Most other animal related idiographic research has been done in the clinical setting of animal behavioural counselling (Overall, 1994, 1997).

In the following sections we will address two groups of issues: what is an operational definition of personality, and what constitutes personality and how we explain it; from both a trait and a phenomenological perspective.

3.2. Trait theories of personality

Adjectives like "shy", "bold" or "interested" are part of our everyday language and we use them to describe other people or animals. The ability of languages to produce those verbal descriptors was the starting point for trait related scientific methods using factorial analysis (Eysenck, 1952; Mathews and Deary, 1998).

The use of personality traits has been reported in antiquity with Galen's idea of four temperament types still used today: melancholic, choleric, phlegmatic and sanguine (Matthews and Deary, 1998). Scientific study of personality however started only at the beginning of the last century. The beginnings can be located in 1915 when Webb did his study on "mental qualities". His data are comparable with contemporary personality dimensions (Deary, 1996) and Allport's lexical hypothesis which defines language as an important encoder of personality traits (Goldberg, 1993).

The basic assumption behind trait theories is that we all have dispositions or traits to behave in a certain way in different situations (Pervin and John, 1997). It is all about the likelihood that a person will behave, feel and think in a certain way. If somebody has more pronounced traits that are associated with extroversion, e.g. being sociable and outgoing, the prediction is that those traits will be present across different situations and the person will be described as high on the extroversion dimension. One of the main critiques of cross-situational stability of traits, presented by Mischel (1976) elaborates the possibility for an introverted person to become more extroverted in certain situations. Trait psychologists argue that that may well happen, but that a person high on extroversion will be even more extroverted in the same situation (Matthews and Deary, 1998). Mischel's dispute is elaborated by social psychologists (Hampson, 1995) and by the idea of constructivism. According to this personality is not in an internal capacity. It is construed in an intersubjective space between people or people and animals (Crossley 1996). The same line of argument is present in discussing all dimensions in cross-situational stability of traits. A person scoring high on the neuroticism dimension can be anxious or nervous in most situations but not in all. However trait psychologists argue that individuals will vary on those dimensions, which are essentially bipolar, e.g. sociable would have positive loading on extroversion, while unsociable would be on the other pole. Not only people, but animals vary on those dimensions, which we will discuss in the next chapter. Most trait theories adopt the concept of hierarchical organization of traits. Dimensions (Pervin and John, 1997), supertraits (Costa and McCrae, 1995) or higher order factors or superfactors (Eysenck, 1952) all represent higher level factors that are fairly robust and have broad predispositions to being manifested across different situations (Pervin and John, 1997).

Apart from situational and hierarchical issues, trait theories address questions of how to measure those traits and dimensions and what constitutes them, whether biology or social environment. Reliable and valid instruments of measurement are therefore an important part of trait theories and each model has developed its own instrument, mostly in the form of self-report questionnaires or target rating scales (Costa and McCrae, 1992). In this project we use both.

3.2.1. The importance of trait theories for this research

Most animal personality studies use personality traits. They are either borrowed from human models or chosen by behavioural experts for that particular species (Gold & Maple, 1994). Very rarely are the observers themselves asked to provide descriptors, which denominate animals' qualitative aspect of behaviour, either on the level of personality (Dutton et al., 1997) or emotions (Wemelsfelder et al., 2000). Often traits are not factor analysed due to the low number of animals, but grouped in the opposite poles of a "dimension" which remains on a descriptive level.

Animal personality studies are scattered across a wide range of disciplines and there has been an attempt to factor analyse all the traits and group them in dimensions according to well established 3 and 5 dimensional trait models (Gosling, 2001). Approaching animal personality from a trait perspective we are faced with all the issues, problems and components that are present in human trait models of personality. To cross validate animal personality dimensions, we can use one of the well established three or five factor human questionnaires (Morris et al, 2002). Another important contribution of trait theories is the emphasis that is put on the contribution of both biology and heredity, and of environmental factors to the development of one's personality (Pervin & John, 1997).

3.2.2. Trait models of personality

Descriptive personality terms have been an integral part of all trait theories. The idea behind these is quite simple and closely related to our use of language: it is assumed that the most important traits will be encoded in a single term, while those not that relevant will remain on a descriptive level. This assumption was formalized in a lexical hypothesis by Allport and Odbert in the early thirties (see Goldberg, 1993). The hypothesis contributed to an extensive development of trait theories of personality. It has been proposed that human personality can be described by a limited number of trait descriptors that can be collapsed into a smaller number of personality dimensions. Since the lexical hypothesis there have been a number of different models proposing various numbers of traits and dimensions. We will however limit ourselves to the most relevant: idiosyncratic trait model (Allport in Pervin & John, 1997), three dimensional model (Eysenck, 1952), biological perspective on traits (Gray et al., 1991; Zuckerman, 1991, 1992) and finally, five factor model of personality (McCrae & Costa, 2004). The last has been used widely in human personality research, with a few attempts to apply it to animals (Dutton et al., 1997; Gosling, 2001; Morris et al, 2002).

The aspiration of each trait theory we have mentioned is to build up a model with a reliable and valid instrument of assessment, most likely in the form of a questionnaire. Each of them will address the following questions: which are the relevant traits? How do they group in a dimension? and finally How do we define this dimension? Another important question is whether traits have a biological, cognitive or social basis (Pervin & John, 1997; Mathews & Deary, 1998). We now know that all components are important. However, the degree to which each factor contributes to forming personality still remains a subject of research.

All those issues mentioned above are present in animal personality research as well. They may not be addressed as consistently, but in order to study individual differences in animals we need to do that. Therefore the importance of trait psychology is not only to discover personality traits and dimensions in animals, but also to build up a comprehensive theory of animal personality. As with human personality there will be different approaches, so the review in this chapter will include the development of animal personality theories.

3.2.2.1. Allport and classification systems

One of Allport's major contributions to developing the concept of traits was his attempt to define and classify terminology (Pervin and John, 1997). The differentiation between traits as units of personality, and states and activities has been accepted by most trait models, including the contemporary five factor model (Eysenck & Eysenck, 1980, 1991; Goldberg, 1993). Chaplin et al. (1988) recapped the importance of differentiation between traits and states. They argue that traits have time and situation related attributes. They have temporal stability, duration and frequency of occurrence. Situational attribute is consistency over different situations Mischel (1976). States and activities on the other hand are more transient with less stability. Chaplin et al. (1988) give examples of traits, states and activities, e.g. "gentle" is classified as a trait, "happy" as a state and "hesitating" as an activity. Those descriptors come up frequently in animal personality research and making a clear distinction between them is yet a task to be addressed systematically. A number of the human personality related terms are present and animals could be described as gentle by just being seen in a brief interaction with another animal or a human interactor. Therefore in order to assess personality traits the observers have to either know the animals or observe them over time and situations. This issue has been highlighted in our research and will be addressed in the discussion section.

Another classification, perhaps even more present in animal individual differences is that of personality, emotions (Plutchik & Kellerman, 1974, 1980; Plutchik, 1994), behavioural expressions (Wemelsfelder et al., 2000), and styles of interactions (Crossley, 1996; Grajfoner et al., 2002). The latter two are essentially operational definitions of emotions. It has been argued that scientific assessment of emotions in animals is impossible by just observing their behaviour (Broom, 1991a, 1991b), therefore exploring distinctions between personality traits, states, activities and emotions may carry relevance for other areas of animal behaviour research.

Allport also makes a distinction between cardinal, central traits and secondary dispositions (Pervin & John, 1997). Cardinal traits are present very strongly in a person and underlie most behaviour. Central traits are present in fewer situations and secondary dispositions are the least permeable and can be displayed in some situations, but not in others. Allport's emphasis on the idiographic approach or uniqueness of traits present in each individual brings us close to the phenomenological approach and its proponents, Rogers (1989) and Kelly (1955). Interestingly Allport's concept of traits at different levels is similar to Kelly's hierarchy of bipolar constructs. The basic ones are strongly integrated in one's perceptual field so that they determine how that person perceives the world. Those basic constructs are not questioned and modified often. However, we also have less stable constructs that are easily replaced with new ones if experiences direct so.

3.2.2.2. Eysenck and three super-factors

The main difference between Allport and Eysenck is that the first emphasises the lexical uniqueness of individual traits, and the latter has developed a reductionist theory based on the relationship between personality and physiological functioning of the body. Eysenck's model of personality is three dimensional (Eysenck, 1952). He says that all traits can be factor analysed to three super-factors: neuroticism, the psychotic dimension and extroversion. Neuroticism includes traits that indicate emotional stability or instability, extroversion is represented by traits that indicate sociability, activity and impulsiveness, and psychoticism includes traits like coldness, creativity, insensitivity. His idea of personality had a strong influence on contemporary trait models. In his monograph (1952) he opens the discussion with a clear distinction between common-sense and scientific research; in a similar way that e.g. Matthews and Deary (1998) begin their own contribution to trait models of personality. They both emphasise the difference between scientific method and common sense personality assessment. The common sense psychology according to Eysenck "seeks to understand" (1952, p. 9) the behaviour of others, while science wants to describe personalities. Eysenck's emphasis of description is very much a reflection of other naturalistic sciences at the time. Therefore Eysenck's biological perspective on personality seems quite natural. His explanation of the introversion-extroversion dimension manifests differences in how our neuro-physiological system works. In stressful situations introverts are easily aroused and respond better to punishment, while extroverts respond better to reward and take longer to get aroused, which is the reason for seeking external stimulation. Eysenck's concept of extroversion corresponds to arousal level and arousal response and therefore manifests the dual nature of extroversion: sociability and impulsivity. He proposes that individuals are positioned along those three basic dimensions. He also proposes that there is a strong hereditary component forming those dimensions in individuals. All those propositions have influenced further development in trait psychology.

It is not surprising that the idea of animal personality fits well in Eysenck's model. Behaviourism and learning theories have been predominant in animal behaviour science, therefore could be an option to further understand animal personality. There are animal personality studies that deal with this aspect of personality (Goddard & Beilharz, 1984a, 1984b). A number of studies on animal individual differences do not use the term "personality" to avoid accusations of being anthropomorphic (Miles, 1997). They resign themselves to the use of "temperament" (McCreae et al., 2000).

3.2.2.3. Biological trait theories of personality

Animal behaviour research has predominantly used neuro-physiology and quantitative ethograms, therefore biological trait models could be more acceptable for animal scientists. Even though the biology of traits is not the scope of this study, we will give a brief overview of its developments. Eysenck's three dimensional model connects extroversion to the physiology of arousal, but fails to provide a clear physiological explanation for neuroticism and the psychotic dimension or for neurophysiology (Matthews & Gilliland, 1999). Gray et al (1991) on the other hand rotate and rename neuroticism and extroversion as "anxiety" and "impulsivity". The two dimensions are associated with the core in the septo-hippocampal system. Impulsivity relates to the mesolimbic dopamine system and is supported by the behavioural activations system. Psycho-physiological methods, for example electroencephalogram for recording the electro activities of the central nervous system, or methods of recording electro-dermal activities in the peripheral nervous system have been used to associate neuro-physiological activities with the personality dimension. However, a number of problems have been linked with those methods (Matthews & Gilliland, 1999). The most problematic aspects are isolation of specific neural activity and its association with specific psycho-physiological function later translated into a trait or personality dimension. As Zuckerman (1991, 1992) points out, the psychobiology of personality is probably more complex and the methods are yet to be improved. Genetics and the evolutionary importance of personality are significant components of a biological approach. Animal behaviour is an important part of evolutionary theory, therefore the evolutionary value of personality traits and emotions (Darwin, 1998) in hominids should encourage study of animal personality (King et al., 1999). However, "animal personality" is terminology most animal behaviour scientists would use reluctantly, as presumably being anthropomorphic. Progressive development in science and research methods is reflected in Zuckerman's statement (1991): "Psychology is not the only science that makes use of hypothetical constructs such as traits. Until 1950s the gene was a hypothetical construct in the science of genetics." (pp. 89).

While Eysenck's superfactor model has a strong biological platform, the five factor model still associated with Eysenck's model, departs from biological definition of traits. The five factor model of personality is more phenotypical than reductionist. It may be seen as a compromise between Allport and Eysenck. The relevance of the five factor model for animal personality research lies in the ease of its adaptability. In his extensive animal personality review, Gosling (2001) meta-analysed aver 200 animal personality studies and produced a list of cross-species personality dimensions based on the five factor model. In our study we refer to the same personality model when identifying personality dimensions in horses.

3.2.2.4. The five factor model of personality

The five factor model of personality was pre-empted by Norman and Goldberg who both found five basic personality factors that were consistent (see Pervin & John, 1997; Matthews & Deary, 1998). There is a strong influence of Allport's lexical hypothesis and Eysenck's three dimensional model (Matthews & Deary, 1998) and it parallels an early six factor model (Deary, 1996). The model was further developed by Costa and McCrae (1987), who particularly contributed to factor definitions. Today the idea of five personality dimensions is integrated with most trait models. The five dimensions are Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism. They consist of facets that are based on adjectives. Proponents argue similarly to Eysenck, that all dimensions are present in every individual, who can score high or low on each dimension. As most other personality models, the five factor model has its own instrument of assessment, the NEO Five factor inventory (Costa & McCrae, 1992). Most instruments for personality trait assessment are in the form of self-report or target-rating questionnaires. Costa and McCrae have developed both methods and present correlations between results from self- and target-ratings (Costa & McCrae, 1992). Like Funder (1995, 1997), Kenny and Albright (1987) and Kenny et al (1994), Costa and McCrae also report a high correlation between self- and target- rating. This issue is particularly relevant for our research, as animal personality assessment depends on target-rating. In order to shift animals conceptually to the target-level we need to change the perspective and reformulate animal behavioural models (see Chapters 5 and 6 for further discussion). The five factor model appears to be consistent not only across human languages and cultures, but also in cross species settings (Gosling 2001; Morris et al, 2002).

The beginnings of animal trait models of personality resemble the early days of human trait development. As with human personality models, most animal studies start off with a number of adjectives used to describe the animals in question. As early trait psychologists collected potential personality descriptors, so do most animal researchers. However, with the exception of Gosling (2001), nobody has consistently systematised and cross validated animal personality descriptors. Therefore most studies either borrow adjectives from previous research or generate their own. In our project we use repertory grid technique (Kelly, 1955) as a method for data collection. We also use further analysis to identify horse personality dimensions.

3.2.3. Benefits of trait personality models for studying animal personality

One of the important questions is what forms personality. It is widely accepted that in human trait personality both genes and social environment contribute to the development and retaining of personality traits. One of the effects of studying animal behaviour is to increase the level of explanation and understanding of human psychology. There has been an increase in genetic research of human personality traits, focusing on monozygotic and dizygotic twins (Spinath et al., 1999). One of the outcomes of studying animal personality is that we have more opportunities to investigate the behaviour and personality of animal twins (Scott, 1987), and therefore gain understanding of how human personality develops and what constitutes it (Gosling, 2001). Ruefenacht et al. (2002) investigated genetic and non-genetic effects on behavioural traits in German Shepherd dogs and reported high genetic correlation with "self-confidence" and "nerve stability". Most of the research however has focused on investigating behavioural traits within species and made comparison between traits in e.g. domestic dogs and humans (Svartberg & Forkman, 2002). As mentioned above, genetic or biological explanation of personality may contribute to an acceptance of personality terminology for animals.

3.3. Phenomenological theories

Phenomenological psychology offers a different perspective on explaining what constitutes one's personality. Most phenomenologists do not really talk about personality as a separate phenomenon. The main idea behind phenomenology is the importance of self-actualisation. According to Rogers (1989) one's personality is less stable and it is moulded by the ongoing experiences that person has. What determines behaviour is not personality traits, but the phenomenal field which is the present reality for that individual and an ultimate source of information. It is a form of self report, similar perhaps to self-report personality questionnaires. Striving for self actualisation is the basic human motivation. The idea of this basic motivation can be extended to animals as well (Wemelsfelder, 1999a).

Kelly (1955) developed his theory of a person as a scientist who anticipates future events by forming and testing predictions that are based on past experiences. Like Rogers, Kelly did not use the term "personality" nor did he develop a personality terminology. An individual is determined by its own construction of reality, which is a consequence of predictions and building up bipolar constructs. Those constructs are the conceptual distinctions an individual makes. In the same way as Eysenck's and Allport's traits, Kelly's constructs can be basic or core and peripheral. Core constructs do not change often, while peripheral, are subject to constant change.

The relevance of phenomenological theories for this project is in their ability to emphasise two things: the interpersonal construction of personality and subjectivity. In the experimental part of our project we used phenomenological principles when collecting the data. We also extended Kelly's idea of a person as a scientist and will be discussing it further in Chapters 5 and 6.

3.4. Comparison of personality models and discussion

There are two issues that need to be highlighted. The first is the fact of conceptual differences between a reductionist trait approach and a holistic phenomenological one. In our opinion both are important and can be integrated. The second is the dependency of trait theories on human language. Animals, with few exceptions, do not have physiological capacities or the need to verbalize and comprehend that form of communication. However, it does not mean they do not have personality, considering its biological and evolutionary value (Buss, 1989).

Different trait theorists have different ideas about how many personality traits are necessary to classify human personality. Eysenck (1952) proposed three major dimensions and Allport 16 or more. Costa and McCrae (1987) focused on Five factor model (Pervin & John, 1997). Phenomenologists, on the other hand, do not treat the dimensional dilemmas as important and focus on exploring, describing and understanding individuals. This main difference is significant. However both classification and description remain important.

There are similarities between trait theories and Kelly's Personal Construct Psychology. Hierarchy is one of them. Eysenck introduced the idea of super-factors as underlying components of personality. They are built from different traits that are again built from both habitual and specific responses. A similar idea is present in Costa and McCrae's work. However Kelly's hierarchical system is different. Of course there are no traits. They are "replaced" with constructs that are also positioned on different levels. Trait personality dimensions and Kelly's constructs are both based on bipolarity. Traits are defined as predispositions or determinants of behaviour. In order to achieve reliability and validity in personality assessment, traits should be stable. Kelly's core constructs are also stable, but they require the construing of more flexible constructs that are frequently questioned and changed. Those constructs, in the same way as traits, and Roger's phenomenal field determine persons' behaviour. Even though the phenomenological idea of constant change is there, Kelly's core constructs being less flexible will determine the behaviour of a person in different situations and across time. The core constructs are those that have been integrated in the concept of self during childhood. Some traits developed in this period may remain consistent. So even though both approaches to explain human behaviour are different, as the trait is more nomothetic and the phenomenological is ideographic, we can still see parallels. One of the other main differences is that most traits models involve biology and social environment as determinants of personality. Kelly, however, does not address the possibility of hereditary and biological influence on forming constructs and on personality.

Even though there are conceptual differences between the two approaches it is not impossible to integrate the two theories and their instruments of data collection. In our project we attempt to do so. The next chapter will review a number of animal personality studies that have adopted the trait concept of personality. Their work started in a similar way to human trait psychology. Adjectives were either collected form human personality studies (Gold & Maple, 1994) or provided by the species experts. Adopting traits as measuring units for animal personality we are, with few exceptions (Pepperberg, 1983), faced with animals' verbal inability. Therefore if traits are based on a lexical hypothesis we have a substantial problem with animals' physiological incapacity to perform self report or projective tests.

To resolve this dilemma, language recognition can be substituted with behaviour. According to phenomenology, behaviour and body movements are, just as words, manifestations of the self (Merleau-Ponty, 1963; Hacker, 1997; Priest, 1998). Therefore self-report tests and target-ratings can be based not only on verbal but also on behavioural manifestation of one's personality. This change of perspective has been incorporated into our experimental design.

The purpose of assessing personality is to predict subsequent behaviour and performance in various environments, which is an important survival ability for both animals and humans. If human language can express adjectives, grouped into five or three dimensions, behaviour can do the same. The limitations imposed by human language should not be reflected in limitation of personality. If humans can rate animal personalities with the purpose of predicting subsequent behaviour, we should assume animals can do the same. We believe that in an animal context the lexical hypothesis leaves us without an instrument of assessment. These conceptual limitations of trait theories suggested considering phenomenological theories as well.

To sum up, in this chapter we have looked primarily at the trait personality theories most relevant for our project. We have also introduced phenomenological models and compared the two concepts. Phenomenological theories will be further addressed in Chapters 5 and 6. The purpose of this chapter is also to provide a theoretical framework for the next chapter, where we will review animal, and specifically, horse personality research.

CHAPTER 4

Animal Personality Review

4.1. Introduction

In this chapter we will first look at the dynamics in the development of animal personality research, discuss its benefits and finally review the relevant studies. The review will highlight theoretical and methodological aspects. In the first part of the chapter we will focus on personality in primates, farm and companion animals; and in the second on horses. At the end we will summarise the chapter and specify our hypotheses.

The beginnings of human personality science were influenced by the lexical hypothesis (Goldberg, 1990) and biological explanation of personality (Eysenck, 1952 Gray et al., 1991; Zuckerman, 1992; King et al., 1999), discussed in the previous chapter. Eysenck's biologically based trait model has been further developed and integrated in various trait personality theories, including the five factor model (Costa & McCrae, 2004). Simultaneously clinically oriented, idiographic personality theories developed (Kelly, 1955; Rychlak, 1981; Rogers, 1989). Their purpose was to understand the individuals, their relation to others and deviations in their behaviour and mental processes (Pervin & John, 1997).

Animals have often been used as models in many areas of psychology: developmental psychology and attachment, the study of emotions and brain functioning, learning and memory (LeDoux, 1998). The psychology of individual differences, on the other hand, has not really included animals. The use of animals as human models encouraged the development of animal behaviour science as a separate discipline. Behavioural models were based on the biology of behaviour and the stimulus-response principles of behaviourism. The goal of studying animals was to identify and explain the mechanisms of behaviour. It was assumed animals would all respond to stimuli in a similar way, following the same behavioural patterns. It is not therefore surprising that individual differences were regarded as undesirable noise and ignored or eliminated in the process of data analysis.

The same field and lab researchers spontaneously used qualitative descriptors of behaviour to identify the individual animals (Hebb, 1946; Goodall, 1999). Combining those descriptors and using them to communicate the information about individual animals integrated the information and in fact created their personality profiles. These profiles were informally recorded as anecdotes with no scientific credibility. Recently those descriptions have been studied systematically and their reliability and validity have been tested (Gosling, 2001). However, animal personality research has developed slowly and has theoretically and methodologically lagged behind the study of human individual differences. The reasons are most likely ontological and social. Questions about what constitutes humans and separates us from other animals are essential to human identity. Our society is primarily a human society. In this social and legal structure animals are not defined as "subjects of life" (Regan, 1984) and therefore their behavioural differences do not have the same impact on the dynamic of the social structure.

Assessment of personality is closely linked to the use of language and to the ability to express verbally the most important traits. Animals' inability to use human spoken language has most likely contributed to the underdevelopment of the study of animal personality. Language has been conceptualised in different ways, for example

Chomsky (1972, 1986) defined the ability of language as an exclusively human capacity acquired through a mutation. The impact of this combined with the Lexical hypothesis categorically excludes all other species, regardless of the research in language communication in primates and parrots (see Pepperberg, 1983; Terrace, 1985; Fouts & Fouts, 1993).

A number of researchers have addressed the question of the evolutionary value of personality (Slater, 1981; King et al., 1999). The two most proclaimed human personality dimensions, neuroticism and extroversion (Eysenck, 1952; Costa & McCrae, 2004), have been consistently identified in most animal species (Gosling and John, 1999; Gosling, 2001). Investigating these dimensions and their association with performance, health and welfare, can contribute to our knowledge about the evolutionary role and importance of personality in animals and humans. It also contributes to our understanding of animal behaviour, performance (Goddard & Beilharz, 1983, 1986; Svartberg, 2002) and welfare (Weiss et al., 2006).

4.2. The Development of the research on animal personality

The study of animal individual differences has been approached from different perspectives: Gosling (2001) did a cross species comparison using the five factor personality model, Wemelsfelder (1999a) addressed theoretical concepts and Dutton et al. (1997) and Wemelsfelder et al. (2000) introduced novel methodologies. We will highlight all these aspects. We will also try to identify studies which explore observers' individual differences and their influence on animal personality ratings. Finally we will review research on horse personality or temperament and their impact on horse performance.

Animal personality studies have been scattered around various disciplines, from agriculture to sociology. Gosling (2001) provided a review paper, gathering over 200 studies on animal personality. The most relevant aspect of this paper was the emphasis on psychometric parameters, reliability and validity. Meta-analysis of the results indicated the existence of traits, comparable with the five factor model of personality (Costa & McCrae, 2004).

There are several papers which have had a significant impact on the development of the theory of animal personality and on our project. The first two are papers by Hall (1941) and Hebb (1946). Hall describes the evolutionary importance of individual differences or temperamental traits in animals, defined as "fearfulness", "timidity", "emotionality", and "wildness". He addresses the hereditary value of traits and emphasises that traits like "cooperativeness", "friendliness" and "gregariousness" have not yet been studied in animals (Hall, 1941, p. 940). Hebb (1946) on the other hand summarises the dilemma of lab researchers saying that "…chimpanzees at the Yerkes Laboratories…have stable differences of behaviour which simulate all the varieties of human temperament." (p. 88). He also suggests that "anthropomorphic terminology" (p. 88), even though very undesirable, may have a significant value for describing and classifying behavioural differences in animals. Hebb also argues that in order to identify those traits the judge needs to have some degree of familiarity with the animals concerned.

Methodologically three studies are particularly interesting: a study by Stevenson-Hinde and Zunz (1978), using adjectives based on the behaviour of animals that were being rated, a study by Dutton et al. (1997), using Kelly's repertory grid technique to assess chimpanzee personality, and a study by Wemelsfelder et al. (2000) introducing free choice profiling for assessing behavioural expressions in pigs. All three were idiographic in two different ways: they provided observers' individual constructions of personality traits (Dutton et al., 1997) and individual profiles of pigs' behavioural expressions (Wemelsfelder et al., 2000) and monkeys' personalities (Stevenson-Hinde & Zunz, 1978).

We will first review primate personality studies, then farm and companion animals, and at the end horse personality and temperament studies. All three groups of research are relevant for our project. Research on primates has an important methodological and historical value. Horses are classified as farm and companion animals and the research on horses refers to studies on both.

4.2.1. Primates

The study of individual differences in primates started within comparative psychology, addressing developmental issues (Spencer-Booth & Hinde, 1969; Hinde & Spencer-Booth, 1971), and investigating psychopathic personality (Buirski & Plutchik, 1991; Lilienfeld et al., 1999). Hebb (1946) was one of the first researchers who emphasised the importance of emotions and temperament in chimpanzees, based on behavioural observations. Since then the methodology of assessing individual differences in primates has developed in different directions.

Some researchers focused on neuro-endocrinal functions (Capitanio et al., 2004) and quantitative characteristics of behaviour (Spencer-Booth & Hinde, 1969; Hinde & Spencer-Booth, 1971; Lilienfeld et al., 1999) or temperament (Clarke & Boinski, 1995). Others adopted the Emotional Profile Index (EPI), an instrument, developed by Plutchik and Kellerman (1974) to measure emotional dimensions, e.g. "rejection", and "exploration" (Kellerman & Plutchik, 1968, p. 1113). The instrument refers to a set of adjectives, which could be used to describe either emotions or personality traits (Kellerman & Plutchik, 1968). To assess primate personality, the animals were scored on adjectives from EPI by several judges, who needed to reach a degree of agreement for the assessment to be reliable (Buirski & Plutchik, 1991). Martau et al. (1985) made a distinction between familiar and unfamiliar raters judging four different primate species, and reported significantly higher inter-rater reliability in the group of familiar judges.

Further studies in primate personality departed from EPI. Stevenson-Hinde's and Zunz's (1978) field study assessed the behaviour of Rhesus monkeys. The experimenters composed a scoring list of 33 behaviourally defined adjectives. Those were based on descriptions of monkeys provided by experienced field observers.

They were identified as "confident-fearful", "active-slow" and "sociable-solitary" reportedly stable over three years of assessment (Stevenson-Hinde & Zunz, 1978).. Gold and Maple (1994) assessed the personality of captive gorillas. Twenty-five behaviourally based adjectives were collected from previous studies on gorilla behaviour. The list was provided to the observers who knew the animals well. The dimensions they identified were "extroverted", "dominant", "fearful" and "understanding" (Gold & Maple, 1994).

Rating lists were mainly provided, though some were partially generated by the observers who scored the animals. No distinction between provided and elicited adjectives has been made. Methodologically very different is a study on chimpanzee personality by Dutton et al. (1997). They used a method, Repertory grid technique (see Kelly, 1955; Fransella et al., 2004), which allows the observers to generate their own personality descriptors. The animals were rated, however, the aim of Dutton's work being to gain access to observers' integrative construction of chimpanzees, and to allow the observers fully to express how they saw chimpanzees' behavioural traits. The results showed that the observers generated similar descriptors which were grouped into six factors: "dominance", "sociability", "machiavellianism", "anxiety", "playful" and "food-related" (Dutton et al., 1997).

4.2.2. Other species

Personality traits and specific behaviours have also been investigated in other species. The "shy-bold" dimension was identified in Pumpkinseed Sunfish (Wilson et al., 1993). A principal component analysis revealed that octopuses vary on three dimensions: "activity", "reactivity" and "avoidance" (Mather and Anderson, 1993), and mink were reported to vary on explorative, fearful and aggressive behaviour (Hansen, 1996). The same behavioural traits were reported for Dumpling Squids (Sinn & Moltschaniwskyj, 2005). Those studies used quantitative ethograms or biological characteristics and did not have judges to rate the animals. Some of the results may have been closer to behavioural expressions or emotions than personality (Kellerman and Plutchik, 1968).

In contrast, for the assessment of personality in spotted hyenas Gosling (1998) used a provided list of adjectives and expert observers. The dimensions he identified were "assertiveness", "excitability", "human-directed agreeableness", "sociability" and "curiosity". Dutton and Andersson (2002) conducted a personality study on six captive Royal pythons, rated by their owner, using a combination of adjectives generated by the owner and a set of adjectives from a previously composed rating scale (Gosling & Bronnenburg, 1998). The study reports two dimensions: the first was interpreted as "warm-cold" and the second "shy-bold" (Gosling and Bronnenburg, 1998; Dutton & Andersson, 2002).

4.2.3. Farm and Companion animals

Farm and companion animals have the most frequent contact with humans; therefore their individual differences usually play an important role in our mutual interactions.

The farm animals most studied are pigs, dairy animals and cattle. Most of the pig studies are methodologically similar, using three types of tests: quantitative ethograms and various techniques to sample data by recording animal movements, postures and vocalizations; behavioural tests, most frequently novel object tests, where animals are exposed to new stimuli; and handling response tests, recording animals' response to a handler (Lawrence et al., 1991). As a rule, the term "personality" is consistently replaced by "individual behaviours" or "temperament", and the terminology is kept technical to avoid anthropomorphism. The situation seems similar to what is described by Hebb (1946) in the primate research community. The primary question in these studies is not how many and which personality dimensions animals have, but whether the individual behaviours are consistent. Lawrence et al. (1991) conclude that there are "stable temperamental characteristics" (p. 83) in pigs, but they fail to demonstrate consistency of behaviour over time and different situations. Similarly Jensen et al. (1995) conclude that in comparison to some other species, pigs do not have "consistent individual behavioural strategies", even though principal component analysis suggests three personality traits: "aggression", "sociability" and "exploration" (Forkman et al. 1995). Hessing et al. (1993) on the other hand conclude that consistent behavioural characteristics in pigs exist in social and non-social situations. Spoolder et al. (1996) argue that pigs demonstrate response consistency over a short period of time and in the same context, but fail to do so in different situations and therefore behavioural types of gilts do not exist. Ruis et al. (2000) have joined the debate and by subjecting the pigs to a variety of behavioural and physiological tests, show that "aggression" is a stable trait. Hayne and Gonyou (2003) group behavioural data into "approachability", "exploration" and "aggression" and indicate that regrouping and traumatic experiences modify behavioural reactions, and that inconsistency in behaviour is not a proof that behavioural traits in pigs do not exist.

A completely different theoretical and methodological perspective has been introduced by Wemelsfelder et al. (2000). Like Dutton et al. (1997), Wemelsfelder implemented a new method, free choice profiling, which allowed 18 observers to use their integrative impressions of 20 pigs. The observers generated their own descriptors by observing the pigs in a human interaction test and rated the animals on those descriptors (see Wemelsfelder et al., 2000; Wemelsfelder, 2001). Methodologically and theoretically the study is one of its kind: it adopts phenomenological framework and uses the behavioural style of pigs as an operational definition for behavioural traits. The experimenters ended up with 18 different scoring lists, one per observer. Using generalized procrustes analysis they calculated the degree of agreement between the observers by correlating individual observer plots with general consensus profile. The agreement between the observers was highly significant and the assessment therefore reliable. The study identified two dimensions of pigs' behavioural expressions. However, these were not labelled. The study also provided pig plots, which show variation between pigs along the two axes.

Studies on personality and behavioural differences in dairy animals and cattle have, similarly to the majority of pig research, also been using standardised behavioural tests and physiological measurements. However, Kilgour (1975) asked two dairymen to rank cows on a single dimensional temperament scale. We have identified an interesting connection between temperament ranking and performance: a placid cow was also described as "an ideal type" and flighty cow as "Comes close to being culled as a 'nuisance' cow." (p. 618). The author does not discuss that. The study reports significant correlation between the raters, and emphasises the importance of the "man-cow interaction factor" (ibid, p. 622).

Consistency of behaviour is, as with pigs, important in dairy animals. Lyons et al. (1989) did a study on the temperament of dairy goats. They used standardised behavioural tests and highlighted three points: the paradox of change and consistency in temperament, the importance of familiarity with the goats, and finally the fact that individual differences predict pituitary-adrenal activity. Schrader (2002) also investigated how consistent the behaviour of dairy cows is in their home pen, using behavioural tests and quantitative ethograms of spontaneous behaviour. Consistency was only identified in spontaneous behaviour, which can predict coping capacity in individual cows. When assessing individual differences in cattle, Tulloh (1961) used six temperamental scores to summarise the behaviour of cattle: "docile", "slightly restless", "restless", "nervous", "wild" and "aggressive". Tulloh relates temperament scores with other measurements and reports that more docile animals are heavier, and animals with a "bad temperament" are also difficult to handle. Grignard et al. (2001) investigated how individual cattle respond to handling and reports that there is variability between sires and their off-springs in handling situations, which affect the selection program.

The importance of personality in companion animals reflects their emotional value. As a consequence pets are the only group of animals whose relationship with owners has been consistently addressed (Podberscek & Gosling, 2001). A brief cross species review of personality traits in pets has been provided by Podberscek and Gosling

(2000). In the following section we will focus on a few studies of cats and dogs, as they could be put in the same category of domesticated animals as horses.

Apart from their potential in animal assisted psychotherapy, cats are mainly our companions. As with other companion animals, some personality traits are more desirable than others. Cats' personality is perceived as the most important factor of their interactions with humans (Turner et al., 1986; Bradshaw, 1992; Bradshaw et al., 1996). Turner et al. (1986) reported that inter-observer concordance was highly significant when animals were rated on "friendliness". These scores were highly correlated with quantitative assessment of cats' behaviour towards humans, e.g. sniffing, head and body rubs. An important contribution was made by Feaver et al. (1986) who combined standardised quantitative ethogram with informal qualitative ratings of cats' behavioural style on 18 provided and predefined adjectives. These were adapted from a previous study on primate personality (Stevenson-Hinde and Zunz, 1978). The cats were rated by familiar observers, using visual analog scales, similar to Wemelsfelder et al. (2000). Significant correlations between formal and informal assessment demonstrates the validity of qualitative integrative assessment of cats' behavioural styles. The importance of behavioural differences in social organization in cats (Durr & Smith, 1997) and wolves (Fox, 1972) seems to confirm the evolutionary value of different personality traits, and their stability over time.

Compared to cats, dogs have always been, like horses, both companions and working animals. The relationship between personality and performance is crucial, especially when selecting puppies to be trained as guide dogs (Goddard & Beilharz, 1986). The results from Goddard and Beilharz's study, however, show that variations in fearfulness between puppies at the age of four weeks do not reflect their responses as grown up dogs. The heritability of behavioural traits has been further investigated by Ruefenacht et al. (2002) who indicates low level of improvement in predictability using a standardised behavioural test on German shepherd dogs. Murphy (1995, 1998) also investigates the assessment of temperament categories, e.g. "anxiety", "suspicion", "excitability" etc. of potential guide dogs. These temperament categories were predefined and identified by dogs' trainers. Even though all categories were based on quantitative description of behaviour, different trainers pointed out different traits in dogs. Murphy reported that trainers' personalities were one of the main limitations and difficulties of the study. Five main traits: "playfulness", "curiosity/fearlessness", "chase-proneness", "sociability" and "aggressiveness" were identified over various dog breeds using standardised behavioural tests (Svartberg & Forkman, 2002). Apart from "aggressiveness" all other traits were grouped into the dimension "shy-bold", which was related to dogs' performance. Bolder dogs reportedly performed better in training and behavioural tests (Svartberg, 2002).

Standardised behavioural tests are prevalent in personality or temperament studies of farm and companion animals. The main interest of the researchers is the consistency of individual behaviours and their relation to various types of performance, from milk production to ability to work as guides.

4.2.4. The development of horse personality research

Horses seem to be of one the most versatile species integrated into human society. For centuries they have been an object of art and an invaluable means of transport. They have been described as one of the strongest and most graceful domesticated animals. Their role as companion and farm animals has expanded and humans have used them in sports and as working animals, in horse assisted therapy and as police and military supports.

An instrument to assess horse personality or temperament can be applied as a prognostic indicator in all these situations. Their individual differences are also important for assessing subjective feelings and welfare, and could be applied in equine conservation (Mills, 1998; Boyd & Bandi, 2002). Temperament in donkeys has been investigated with the purpose of finding the best matching foster homes (French, 1993).

Horses' behavioural differences and personality therefore have implications in many areas: heritability and breeding programmes (Worth-Estes, 1952), health and welfare (Waran and Cuddeford, 1995), including abnormal behaviour (McAfee et al., 2002), performance (Visser et al., 2001; Visser et al., 2002), including racing (Hutson & Haskell, 1997), show-jumping (Visser et al., 2003) and in therapeutic riding programs (Potter et al., 1994; Anderson et al., 1999).

General texts about horses (e.g. Hafez, 1969; Fraser, 1992; Kiley-Worthington, 1997; Budiansky, 1998; Mills, 1999) describe mechanisms of behaviour but do not address personality or temperament to any great extent. However, this has been done in an increasing number of research reports on horse behaviour. Most of them can be categorised in four groups:

1) behavioural, physiological, welfare and learning differences using standardised quantitative behavioural and physiological tests (Mal et al., 1991; Waran & Cuddelford, 1995; Winskill et al., 1996; Hanggi, 1999; McAfee et al., 2002; Saslow, 2002; McCall et al. 2003; Murphy et al., 2004; McCall et al., 2006).

2) individual differences in social interactions (Wood-Gush & Galbraith, 1987; Morgan et al., 2000; Boyd & Bandi, 2002) and handling (Mal & McCall, 1996; Sondergaard & Halekoh, 2003; Sondergaard & Ladewig, 2004; Lansade et al., 2004);

3) temperament and personality assessment using quantitative ethograms and standardised behavioural tests (Wolff et al., 1997; Seaman et al., 2002), using rating scales (French, 1993; Mills, 1998; Morris et al., 2002) or combining both (McCann et al., 1988; Le Scolan et al, 1997; Visser et al., 2000; Visser et al., 2001; Momozawa et al., 2003; Sian Lloyd et al., 2006 in press);

4) individual differences and performance in sports (Worth-Estes, 1952; Hutson and Haskell, 1997; Visser et al, 2003; Buckley et al., 2004) and horse assisted therapy (Potter et al., 1994; Anderson et al., 1999).

The first two groups of studies used standardised tests to record quantities of behaviour: postures, vocalization and movements; but also take physical measurements like heart rate (Waran & Cuddelford, 1995). Movements and postures were an operational definition of social interactions between geldings and ponies

(Wood-Gush & Galbaith, 1987), young horses (Sondergaard & Ladewig, 2004) and a free living harem of takhi (Boyld & Houpt, 1994; Boyld & Bandi, 2002). Studies on handling horses emphasise the importance of temperament, particularly fearfulness (Lansade et al., 2004). Fearful horses are more prone to accidents and harder to manage (Mal & McCall, 1996). Sondergaard and Halekoh (2003) reported that frequent handling decreases heart rate when horses are introduced to a novel environment.

Studies listed in the last two categories are methodologically more versatile. They use either behavioural tests or rating scales, similar to human personality questionnaires. Sometimes, a combination of both is applied.

Quantitative measurements are developed within a theoretical framework that uses temperament or behavioural categories. Focus is, as with most farm animals, rather on consistency of behaviour than on identification of personality dimensions (Wolff et al., 1997). Temperament (Mills, 1998; Seaman et al., 2002) is often defined as "biologically rooted individual differences in behaviour tendencies that are present early in life and are relatively stable across various kinds of situations and over the course of time" (Lansade et al., 2004, p. 132). This definition comes from human developmental psychology. It reflects the sort of definitions used in most temperament studies of farm animals. Researchers using quantitative tests label temperament factors as "overtly responsive" and "active"; and "overtly passive" (Seaman et al., 2002). Those using personality adjectives use "obduracy" associated with "difficult to handle" and "spiteful"; and "vivacity" associated with "playful" and "outgoing" (French, 1993).

The further review will focus on horse studies which partially or solely employ the method of rating horses on personality adjectives. We will emphasise three aspects: whether adjectives are provided to the observers, whether the raters are familiar with the animals; and, which personality dimensions they identify.

By defining temperament as an interaction between donkeys and people, French (1993) offered a categorically different definition of temperament. In her study the adjectives were provided to the raters and ordered as bipolar pairs e.g. "calm-nervous". Donkeys from a Donkey Sanctuary were rated by the members of staff and the foster families using a distance scale for each pair of adjectives. The observers had significant inter and intra-observer consistency for the adjectives and the donkeys. This study is important for its methodology, emphasising the donkey-human interactions and the value of successful re-homing of the animals.

Morris et al. (2002) used a modified NEO-PI five factor personality questionnaire to assess 10 horses by nine judges. The raters, familiar with the horses, were given a questionnaire for each horse, and asked to complete it in their own time. Kendall's W showed the highest consistency when scoring horses on items associated with Extroversion, Neuroticism and Agreeableness. The results demonstrate that Costa and McCrae's (1992) questionnaire is a reliable instrument for assessing horse personality (Morris et al., 2002).

Mills (1998) on the other hand reports that the personality adjectives used to describe 20 horses failed to achieve agreement between raters. He gave the adjectives to seven observers, who provided their own definitions. Both coefficients Spearman rho and Kendall's W were calculated. However, only the significance level for Spearman rho was shown and the observers agreed only on two adjectives, "flighty" and "sharp". Mills argues that the adjectives are not consistently defined in an objective way and therefore they cannot be used as a reliable and valid instrument.

To address the question of validity, some researchers combine standardised behavioural tests with observers' rating and correlate the results.

Sian-Lloyd et al. (2006 in press) assessed horses using quantitative ethograms and ratings by three judges. Thirty predefined adjectives were selected from a previous study on primates (Stevenson-Hinde and Zunz, 1978). The results showed high observer agreement and high correlation between the scores provided with both methods. The study demonstrated that handlers, more familiar with the horses, achieved higher degree of agreement. The personality components in this study were "dominance", "anxiousness", "excitability", "protection", "sociability" and "inquisitiveness". The authors offer an alternative interpretation of dimensions by linking "anxiousness" and "excitability" to neuroticism; "protection", "sociability" and "inquisitiveness" to extroversion; and "dominance" to agreeableness (Morris et al., 2002; Sian Lloyd et al., 2006).

McCann et al. (1988) used a one-dimensional approach ("very nervous, nervous, normal, quiet") to address emotionality evaluations. Four experienced horsemen scored yearlings by ticking one of the four degrees of nervousness. As with the previous study, the results were highly correlated with quantitative assessment of behaviour and heart/ respiratory rate. Nervous yearlings moved around more and their heart/ respiratory rate was significantly higher.

Le Scolan et al. (1997) correlated the results from standardised behavioural tests with ratings on seven provided predefined adjectives, all referring to "nervousness" and "sociability", elicited from the riding teachers. The correlations between the two tests were significant; more fearful horses took longer to cross a bridge, which represented a novel object and a handling situation. Similarly, there was a high correlation of scores between temperamental factors "anxiety", "novelty seeking" and "understanding" and responses to both a balloon reactivity test and frequency of heart rate and defecation (Momozawa et al., 2003). More anxious horses were unwilling to touch the balloons and had a higher heart rate and greater frequency of defecation. Visser et al. (2000) subjected 16 adult horses to two behavioural tests and had them rated by riders on 10 temperamental adjectives (e.g. "confident", "brave"). The inter-rater reliability was significant and the two personality factors were "responsiveness to environment" and "attention to the rider". Results from both qualitative and quantitative tests were significantly correlated. Visser et al. (2001), similarly to Mills (1998) propose "objective" definitions of descriptors, which would reduce variation between individual raters and eventually reduce the number of raters needed for reliable assessment of horse personality.

As horses are used for various purposes, their performance is crucial. Choosing a horse that will match the rider's personality will increase the safety and welfare of both the horse and the rider (Visser et al., 2001). The type of horse performance regarded as desirable depends on the work. Sports, in particular show jumping (Visser et al., 2003b), require a horse's ability to learn (Visser et al., 2003a), and personality traits can affect that. Visser et al. (2003a) classified horses as performers (completing the test) or non-performers (not completing the test), using aversive and reward based learning tests. Correlation between general learning tests and show jumping showed that performers were also better in show-jumping.

Methodologically interesting was a study on the temperament of thoroughbred broodmares, and the racing performance of their off-spring (Worth-Estes, 1952). The personality traits collected were based on the behaviour of 150 mares. The mares were rated by experienced horsemen, familiar with the animals. A hundred mares were eventually selected for the study and grouped as either "high-spirited" or "phlegmatic". Worth-Estes' methodology resembles that of Dutton's et al. (1997), Wemelsfelder's et al. (2000) and Grajfoner's et al. (2002). In these studies the observers generated their own descriptors and assessed the animals on them. In spite of the interesting methodology, the results showed no relationship between the mares' temperament and the performance of their off-spring, measured by the amount of money earned (Worth-Estes, 1952). Not even quantitative ethograms describing single behaviours, e.g. movement or posture in horses just before a race, were successful in predicting the finishing order or the winner (Hutson and Haskell, 1997).

Apart from show-jumping and racing, horse assisted therapy is another area where personality can affect horse performance (Anderson et al., 1999). Three experienced riding instructors rated 103 horses on 20 personality traits, adapted from French (1993). Anderson et al. (1999) also measured plasma cortisone, norpinephrine and epinephrine along with behavioural reactions to a reactivity test. The results from the rating showed low score correlations of less than 40%. There was also no correlation between personality, reactivity and hormone concentration (Anderson et al. 1999). However, there was an indication that there could be a relationship between extreme personality traits and quantitative measurements. A considerable proportion of horses with the highest reactivity scores (64%) were in a therapeutic riding program, which shows that the selection of the horses for therapeutic purposes needs to be reconsidered. Highly reactive horses may not be as safe as placid ones.

4.3. Summary and conclusions

In this chapter we have reviewed animal personality studies, with particular attention to horses. Our main interest concerned the methodologies the studies employed and personality traits they identified.

Most animal personality studies have designed their experimental work under the influence of biological theories of personality (Gray 1978; Zuckerman, 1992, Eysenck, 1997). Instruments of assessment include standardised behavioural tests and quantitative ethograms, developed within the realm of behaviourism. Therefore it is not surprising that the term "personality" is mostly replaced by "temperament" (Seaman et al., 2002), due to its assumed implicit anthropomorphism (Hebb, 1946). Temperamental or behavioural factors are usually derived from principal component analysis on quantitative ethograms. The experimenters' choice of labelling normally reflects biological roots e.g. "active-passive".

The second large group of instruments of assessment utilise rating scales consisting of adjectives e. g. "shy-bold". These descriptors signify qualitative aspects of behaviour, an individual style of behaviour which is consistent over time and situations (Pervin and John, 1997). Conceptually this is a significant departure from quantitative behavioural tests. The rating scales resemble human personality questionnaires, developed by trait psychologists. In human and animal research alike, those rating scales are mostly composed by the experimenters. Individual adjectives are often defined and then given to the observers. Most of the adjectives have been collected and applied to various human personality studies (Goldberg, 1990). For animal use, they can be taken from previous studies or borrowed from an existing model (Plutchik, 1994). Rarely are the adjectives actually composed on the basis of the behaviour of the animals that are being rated (Stevenson-Hinde & Zunz, 1978; Dutton et al., 1997; Wemelsfelder et al., 2000). Therefore the observers' input is somehow limited, and species specific behaviour can be missed (Dutton & Andersson, 2002). There is a possibility that the observers themselves would not use the adjectives in the way they are being defined or they will make use of another descriptor which is not included in the scoring list.

A few studies employed a method that allowed the observers to express their integrative impressions of animals. This holistic, phenomenological principle was used by Dutton et al. (1997), Wemelsfelder et al. (2000), Grajfoner et al. (2002) and Dutton and Andersson (2002). With that, they depart further from reductionist methodologies. Another indicator of a holistic or more integrative method of assessment is focus on individual animals and provision of more in-depth information on their personality profiles. Only a small number have reported such profiles of individual Rhesus monkeys (Stevenson-Hinde & Zunz, 1978) pigs (Wemelsfelder et al., 2000; Grajfoner et al., 2002) or dogs (Shapiro, 1990).

Observers' familiarity with animals has regularly been reported, although it has not been systematically investigated. There are no reports of whether familiar observes rate animals differently from unfamiliar ones, and if so what the differences are. It has been suggested, however, that familiarity with animals will increase the accuracy of the assessment (Martau et al., 1985).

There is no uniform classification of temperamental or personality traits, and so most studies use their own descriptors and label the dimensions themselves. Gosling (2001) provided a cross-species review of personality dimensions and labelled them according to the five factor personality model. Most traits reported could indeed be classified as facets of two dimensions, Neuroticism and Extroversion.

Horse personality studies reflect the research on other species: most instruments of the assessment are standardised behavioural tests and quantitative ethograms which are factor analysed. There are a number of horse studies that have used rating scales. The adjectives for those were borrowed from previous studies on other animals (Sian-Lloyd, 2006 in press). None of horse personality studies, with the exception of Worth-Estes' (1952), have actually collected descriptors purely based on horse behaviour. There has not been any investigation as to whether familiarity with horses and / or familiarity with words influences the ratings and affects the reliability of the assessments. Traits have not been classified in a uniform way. Morris et al. (2002) applied an adapted NEO-PI five factor personality questionnaire to horses and reported the highest agreement for Extroversion, Neuroticism and Agreeableness. As horses are used for sports and therapeutic riding, the relationship between personality and performance has been investigated. Even though there seems to be a connection between the two, no definite conclusions have been made.

Throughout animal personality research a fear of anthropomorphism can be detected. Several authors advocate the choice of behavioural tests or predefined standardised rating scales to avoid observers' self projection. However none of the studies empirically investigate whether the observers' individual differences affect their assessment of animals.

In the present study we aim to: i) use a novel application of Repertory grid technique to assess horse personality and investigate its reliability; ii) explore the personality profile of a domesticated horse; and iii) investigate the effect of observers' individual differences on their assessment of horses.

Our specific hypotheses are:

1) Both familiar and unfamiliar observers will provide similar personality adjectives for horses;

2) Familiar observers will have a higher degree of agreement when rating the horses;
3) Observers will use the constructs consistently;
4) Observers will rate more reliably on the adjectives they provide themselves;
5) Horse personality dimensions will be in line with the five factor personality model;
6) Familiar and unfamiliar raters will agree on their horse personality scores;
7) There will be association between horse personality and performance;
8) Observers' personality, empathy and emotional intelligence will not affect their rating of the horses.

We will conclude this chapter with yet another exciting question: Can animals be personality psychologists? This question has been postulated before (King et al., 1999; Morris et al., 2002), but never fully discussed. We will address it in the next chapter along with the introduction of the concept of animal-as-scientist.

CHAPTER 5

Animals as Scientists

5.1. Introduction

Animals are a part of our everyday experiences. We are curious about how they live, so we study their behaviour. We use them for food, work, education and company, and we strive to treat them ethically. We form close relationships with them, and we are concerned about their welfare. They are important to us. We prefer some animals to others. Some of us like dogs, some cats, some horses or parrots. We make preferences based mostly on how animals behave. If we have close interactions with individual animals, there are some we prefer more than others. Some of us like animals that are bold, active and playful; some of us prefer those that are more subdued, shy and calm. When talking about animals close to us we very rarely speak about the quantity of their behaviour: for example how many steps our cat or dog makes in a day. However we can very confidently talk about the behavioural qualities or styles such as how shy or active or nervous that individual animal is (Sanders, 1992).

In our everyday language individual differences in animal behaviour or their personality are nothing unusual. It is common sense that not all cats, dogs or horses are the same, and that there are differences not only between species, but also between individuals. We know that what is important to us in our interactions with animals is not how many steps animals make when they approach us, but what is the style of their approach and interaction. We can tell whether an animal is scared or nervous or friendly. On the other hand the same animal seems to recognise our style of behaving and respond to it. This mutual recognition and consequential interaction is based on the style or quality of individual behaviour. The qualitative aspect of behaviour or qualitative approach to studying animal behaviour is an important concept in this chapter. In everyday life we normally do not question our qualitative assessment of animal behaviour. It makes sense to us, it helps us to understand and interact with our pets, so we use it, almost instinctively.

In science as well as in everyday life we strive to make sense of events, ourselves, our environment and other human and non-human animals. Even if we are not animal behaviour scientists we still build anticipations or expectations of how an animal will behave in a given situation and interact with them accordingly. Our expectations are constantly tested and we either retain or replace the theories or constructs we have. This is in line with Kelly's Personal Construct Theory (1955) and the concept of a person-as-scientist. Theories are our construction systems, hypotheses are our expectations or anticipations and we test them out conducting experiments or experiencing situations and events around us.

5.2. Animals as scientists

How about animals? Do they, like us, construe events in their life and try to make sense of others' behaviour? Do they have expectations of how another individual behaves and what their personality is? If they do, are those expectations based on their previous experiences? Do they experience events, test expectations and alter their constructions?

In this chapter we will present our argument that Kelly's Personal Construct theory (1955) and his concept of person-as-scientist can be adjusted to interpret animal behaviour. We will argue that animals, like humans, try to make sense of their social and physical environment and build up expectations, anticipations or hypotheses, which are subjected to further experiences or tests. As a result, expectations or hypotheses are either confirmed or modified. This cycle of anticipation, experience and adjustment is based on the integration of knowledge, emotions, awareness and personality. The constructs that we generate to make sense of the world around emerge at different times throughout one's life. Therefore one ends up with constructs that may not have a verbal label, but are an essential building block in one's construction system. Animals, like humans, are not necessarily aware of the full structural system of their constructs. Cognitive awareness of nonverbal and pre-verbal constructs like "warmth", "food", "sleep", "comfort" for example, may be low, and these basic constructs may be more efficiently communicated by behaviour, not words (Kelly, 1955).

As few humans, or other animals live in isolation, the process of construing, anticipating and testing is an essential part of our interactions with others and crucial for successful integration and survival. The idea of animal-as-scientist can be used to construct an integrative model of animal behaviour, incorporating both quantitative and qualitative aspects.

In the following sections of this chapter we will first present our theory based on Kelly's Personal Construct Psychology, then look at the theoretical background and focus on intersubjectivism. Finally, we will question whether integrative models of animal behaviour, including the one we propose, are or are not anthropomorphic. To address animals and humans, we will also use the terms "human and non-human animals", and "humans and other animals".

We are not the first to claim that animals make hypotheses. As a part of his learning theory, Krechevsky (1975) did a number of experiments with rats in a multiple-unit discrimination box (see Witkin, 1940; Levine, 1959). He argues that animals are individuals (Krechevsky, 1975, p. 14): the rats adopted one of four different strategies to find a solution to the problem: they tried going right, left, persevere or alternate, and eventually chose the correct way. Krechevsky argues that this is an example of "organized and systematic" (ibid., p. 15) responses and not trial and error behaviour. Animals test behaviour and give it up if they experience failure. Subsequent behaviour will emerge because earlier attempts have not been successful. Krechevsky explains the idea of animal hypothesis further by comparing it to the idea of "means-end-readiness", developed by Tolman (Tolman & Krechevsky, 1933). Tolman and Krechevsky both emphasise animals' selective response to given stimuli; their constant display of systematic behaviour during a learning period; and finally the "self initiating" property of animals' behaviour (Tolman & Krechevsky, 1933, p.62), resulting from previous experiences. Hypotheses in animals should be used to "designate systematic, docile, selective, self-initiated…" behaviour (ibid., p.69). Therefore animals in discrimination learning individually create hypotheses that are tested by systematically trying out different behaviours. The behaviours are individual, goal oriented, and based on previous experience.

Krechevsky's concept resembles the one we have suggested, based on Kelly's Personal Construct Theory (Kelly, 1955). However, there are some theoretical differences: while Krechevsky focuses on hypotheses in learning, expressed in rats' "means-end-readiness" behaviour, Kelly argues that constructs are not necessarily connected with decision making. They are an integration of the totality of our experience in the world. This integration includes emotions, cognition, awareness and personality and therefore avoids fragmentation of any sort. This use of the totality of human experiences when assessing animal behaviour is investigated by Dutton et al. (1997) and Grajfoner et al. (2002). Both studies take on Kelly's Personal construct theory to discover human construction of personality dimensions in chimpanzees (Dutton et al., 1997) and pigs' styles of interacting (Grajfoner et al., 2002). The second study also deals with inter-judgement agreement. The assessment of animals is based on their interactions with their physical and social environment. Animals' consistent behavioural differences and styles of interaction are manifestation of animals' ability to individually make sense of the world by construing it and anticipating future events according to their individual construction systems. Animals' behavioural differences are therefore a manifestation of their individual anticipations or expectations of future events.

Kelly's theory is based on axioms, formulated into the *fundamental postulate* and 11 *corollaries*, which will be fully described in the next chapter. Here we will discuss the corollaries most relevant for explaining the concept of animal-as-scientist.

The fundamental postulate "A person's processes are psychologically channelized by the ways in which he anticipates events." (Kelly, 1963, p. 46) manifests the essential phenomenological idea that a person is inherently active, not reactive, and spontaneously initiates interactions with the environment.

For our purpose we have divided the most relevant corollaries into two groups: those which focus on the individual and their sense making of the world; and those which explain the reality the individual shares with others.

The construction corollary states that an individual "anticipates events by construing their replications." (Kelly, 1963, p. 50). Animals, like humans, construe their anticipations based on their previous experiences. A dog used to going for a walk every evening will anticipate that event and behave accordingly. *The individuality corollary* explains that each individual will construe an event differently, e.g., if there are two dogs in the same household, they will construe going out for a walk individually (ibid, p. 55). One of the dogs may be afraid of cars, so she may construe going out as a traumatic event, whereas the other may construe it as an exciting opportunity to find additional food. *The experience corollary* states that "A person's construction system varies as he successively construes the replication of events." (ibid., p. 72). The owner may change walking location and take the dog to the countryside instead. This will be an unexpected event, and the dog's construction and consequently anticipation or hypothesis of going out may change. *The dichotomy corollary* refers to the assumption that we all have a definite number of bipolar constructs. The constructs can be as basic as "pleasant – unpleasant", "tasty – not tasty", "dangerous – safe", "warm-cold" etc. Constructs can be nonverbal or pre-verbal, to do either with psychology or "covering the realm of physiology…such…as digestion, glandular secretion and so on" (ibid., p. 51). The constructs an animal makes are usually conceptual distinctions such as: "alone-not alone", "no food-food" or "active-passive".

The second group of corollaries concerns interactions with others. They are particularly relevant when we reciprocally construe each others' behaviour in social contacts. Firstly *the commonality corollary* says, "To the extent that one person employs a construction of experience which is similar to that employed by another, his psychological processes are similar to the other person." (ibid., p. 90). If both dogs from the previous example construe the experience of going for a walk in a similar way, "their behaviour will exhibit similar characteristics." (ibid., p. 92). Secondly, *the sociality corollary* states "To the extent that one person construes the construction processes of another, he may play a role in a social process involving the other" (ibid., p. 95). This corollary explains the importance of construing "the other person's outlook" (ibid., p. 95) for any social interaction to be successful. This is important in both ways, humans assessing animal personality and vice versa. Predicting the behaviour of others is an important social and biological act within and between species (Nagel, 1986; Shapiro, 1990). Griffin (1992, p. 64) describes a group of lionesses, involved in the coordinated hunt of a wildebeest. They did not just predict each other's behaviour; they also predicted or construed the wildebeest's construction of the experience, which resulted in a successful hunt, with minimal energy loss. Therefore predicting each others' behaviour and sharing the construction of an experience has a significant survival value. The ability of animals to assess and predict consistent human behaviour or personality is particularly important for the wellbeing of companion and farm animals.

The idea of animal-as-scientist is essentially phenomenological and intersubjective. First, because the process of construction, although idiosyncratic, overlaps with the construction process of other individuals, which is manifested in meaningful interactions; secondly, the individual differences in construction are manifested in individual styles of behaviour; and finally animals' construction is a holistic process that integrates emotion, cognition, awareness and personality. In the next section we will further explain the theoretical framework of our idea and compare it with rationalism, the dominant theoretical framework in animal behaviour science. We shall first depict the philosophical differences between rationalism and phenomenology and then explain the basic premises of intersubjectivism.

5.3. Theoretical background

The predominant scientific attitude to animal emotions, consciousness and to some degree personality as we saw in the previous chapter, is that if it exists, it is internal, and not available for objective scientific scrutiny. This originates in the dualistic philosophy of Descartes (1989) and the rationalists of Port-Royal, which have been an integral part of contemporary mechanistic and rationalistic perspectives in science (see Wemelsfelder, 1999a; Dutton and Williams, 2004). In dualism the inner - subjective is epistemologically independent from the outer - bodily behaviour. Language is the most important means of communication about the internal or subjective experiences. As a consequence, animals, lacking the use of verbal language, are frequently classified as mechanical bodies, which are inert and need something to get started (for further review see Wemelsfelder, 1985; 1997a; Gaita, 2003). This principle is evident in some animal personality studies using quantitative measurements of behaviour (e.g. Lawrence et al, 1991; Jensen et al., 1995).

Both existentialism and phenomenology developed as a response to and critique of dualism and a mechanistic perspective on individuals (Hribar, 1993). According to Wittgenstein (Hacker, 1997) a dualistic split is the root of philosophical error. Existentialism defines individuals, including animals, as active agents, constantly evaluating their position in the interactive environment. Phenomenology focuses on individuals' perception of the situation and self actualisation. There is no mind-body split and subjective experiences are accessible to others through overt behaviour.

Merleau-Ponty formalizes that by introducing the concept of body-subject (Merleau-Ponty, 1963). For a body-subject mental manifestations take the form of expressive, meaningful behaviours which are communicable to others. Individuals are their bodies and experiences are based on their active corporeal involvement in the world (Crossley, 1996). This involvement has a situated corporeal attitude, a way of relating to others, and it is manifested in everything an individual does. This is an intersubjective definition of emotion (Crossley, 1996, Grajfoner et al., 2002). By analogy, an intersubjective definition of personality could mean a more consistent, situated, corporeal attitude or more stable behavioural styles.

Intersubjectivity, developed from existentialism and phenomenology, helps to address the question, whether animals also assess our personality. From the intersubjective perspective, human-animal personality assessment is reciprocal, manifesting a mutually shared construction of reality (Schutz, 1967; Sanders, 1993) or of the sociality corollary (Kelly, 1955).

Crossley (1996) presents two views of intersubjectivism, with their own definition of subjectivity. Firstly, egological intersubjectivism, based on Husserl's idea of a subject, constituted through observation, perception and imagination (Husserl, 1970); and secondly, radical intersubjectivism, which defines subjectivity as an essential constituent of intersubjectivity - an irreducible phenomenon. We will concentrate on the latter, which incorporates the work of Merleau-Ponty and Wittgenstein. In contrast to dualism, in this the body ceases to be a mechanical object, and is restructured as a body subject (Merleau-Ponty, 1963), with an active corporeal involvement in interactions. Mental events which were hidden by dualists, are now inseparable from behaviour and always intersubjectively available. Actions and experiences arise from dialogically constituted situations which are irreducible to individual subjects. Emotions and personality traits in the form of behavioural styles are intersubjectively available, to both humans and animals. Agreement, inherent to intersubjectivity, is achieved through spoken language or behaviour. They are both intelligent forms of communication (Merleau-Ponty.1963), as verbal expressions are only a refinement of naturally expressive behaviours (Wittgenstein in Hacker, 1997).

5.4. Intersubjectivity and animals (non-verbal others)

The importance of expressive behaviours in intersubjective space is demonstrated in the research of pre- or non-verbal subjects: infants and animals. Trevarthen (1985; 1986) reports a series of results from studies on emotional communication between infants and their mothers, using a non-reductionist approach in analysing the data. The work emphasises the evolutionary and physiological value of emotions and consciousness (Trevarthen, 1979; Trevarthen & Aitken, 1994), the development of syncronised movements (Trevathen, 1984) and the musicality of infant-mother communication (Papaeliou & Trevarthen, 1994). Theoretically Trevarthen joins phenomenologists in their emphasis on a subject's inherent activity and dynamic interconnectedness with other subjects and finally in their criticism of Descartes' dualism (Trevarten, 1993a). He also argues that consciousness and cognition are not necessary prerequisites for the existence of intersubjectivity (Trevarthen, 1993b) and that newborn babies are ready to enter an intersubjective space with their mother immediately after birth (Trevarthen, 2001). The evolutionary and survival value of animal semiotics, manifested in behaviour and "a level of intersubjectivity" (Trevarthen, 1990, p. 692) has also been mentioned.

Animal involvement in intersubjective space has been investigated from sociological and psychological perspectives, among others.

In intersubjective terms the social world is not an object; it is a between world generated through the interactions of its members (Schutz, 1967), including animals and humans (Dutton, 2006a, in press). This sociological definition of intersubjectivity is based on shared social resources, common sense knowledge and the individual's location amongst other situated embodied beings, animal and human (Sanders & Arluke, 1993). Social behaviour is reciprocated as both parties are oriented towards each other. For the reciprocity of perspectives to be successful a collective sense making process, based on common-sense knowledge, is essential. Animals are equally involved in this process, as we will see from the examples below. Therefore, animals must have the ability to construe and hypothesise about the behaviour of other animals and humans.

Shapiro (1990) reports a case study based on prolonged interactions between a dog and a human. His theoretical framework is that of phenomenology and radical intersubjectivism, defining the dog as a body-subject with intentional behaviours (p. 192). The outcome implies a mutual understanding between the human and the dog, based on body movements, initiated and directed by both interactors. This mutual understanding through body movements is labelled "kinaesthetic empathy" (ibid., p. 193). Shapiro describes social construction as a set of social beliefs and attitudes (ibid., p. 193) which include both human and non human animals alike. The study reports mutual assessment of behaviour and personality between the dog and the human.

Sanders (1992) conducted a number of case studies, interviewing dog owners about different aspects of interactions with their dogs. He defines intersubjectivity as "two actors, reciprocally oriented toward each other" (ibid., p. 2). He also refers to Schutz's concept of social exchange, which is possible due to a mutually shared semiotic system, whether language or behaviour. Merleau-Ponty, Wittgenstein and Kelly have all emphasised this equal importance of spoken language and behaviour on which our idea of animal-as-scientist is based. Sanders concludes that mutual awareness and engagement in communication that is conventional and agreed on is present not only in humans, but also between animals, and between humans and other animals (Sanders, 1992). Humans do not interact with animals as objects, but with animals as subjects and vice versa. Consequently, an animal will establish a different intersubjective space with a different person (Sanders, 1992; Hearne, 1994, 1995). Using the same analogy, different animals will display their own style of behaving with a person and consequently will create different styles of interaction (Wemelsfelder et al., 2000). This axiom is incorporated in the methodology used by Wemelsfelder et al. to assess pigs' differences in behaviour (2000). Animal attentional styles (Wemelsfelder, 1997a) are defined as ways animals pay attention to an interactive situation. Animals are actively engaged in interactions and co-create their dynamics. These styles can be denominated as "enthusiasm", "fear" etc. Attentional styles were later presented as "behavioural expressions" (Wemelsfelder et al. 2000, p 195) and defined as directly observable expressive aspects of behaviour. It seems that conceptually behavioural expressions correspond with "style of behaviour", which is an intersubjective operational definition of emotions (see Crossley, 1996). The experiment involved a group of human raters, who observed short interactions between individual pigs and a human interactor. The interactor's behaviour was partially standardised, by responding to pig's initiatives rather than taking a lead herself (Wemelsfelder et al., 2000). The observers described the pigs' behavioural expressions using individually generated adjectives. The study reports significant agreement between the observers. According to the authors, this indicates that the observers systematically applied their parameters for assessing qualitative aspects of behaviour. Schutz would define this as common-sense knowledge and a mutually shared semiotic system, which is expressive behaviour (Sanders, 1993). By using this design Wemelsfelder makes the most radical theoretical and methodological jump from the mainstream experimental design in the science of animal behaviour and welfare.

If the condition for successful interactions and mutual recognition of the meaning of behaviour are common-sense knowledge and a shared semiotic system, then knowing the species and individual animals may increase the accuracy of assessment and prediction of behaviour. In human-animal interaction this goes both ways. In animals as well as humans, common sense knowledge and a shared semiotic system influence the experience of an interaction, and therefore affect the anticipations or hypotheses. So far we have demonstrated that in order to have meaningful interactions, animals must be able to make sense of their social and physical environment. They must be able to construe others' behaviour, make hypotheses or have anticipations or expectations and test them out. The animal-as-scientist learns from replications of similar situations by adding new constructs or adjusting already existing ones. This process involves an integration of the totality of an animal as an individual, therefore it is a holistic rather than a fragmented act. As proposing a behavioural model which defines animals as scientists may seem anthropomorphic at first, we will briefly define anthropomorphism and describe its position regarding cognition and consciousness – two concepts that cannot be avoided when discussing intersubjectivity

5.5. Intersubjectivity and anthropomorphism

Anthropomorphism is "a manner of representation, entailing the figurative, erroneous, or naïve attribution of human experiences to animals." (Crist, 2000, p. 7). It can be either an arrogant attempt to explain animal behaviour without considering the animals' own perspective or complete refusal to see any similarities between humans and animals. Both are equally erroneous (Dawkins, 1993; Wemelsfelder, 1999b) and undermine "the credibility, or realistic force, of accounts that in some way picture animal life and human affairs as permeable to one another." (Crist, 2000, p. 7).

We will discuss anthropomorphism in our discussion of the intersubjective model of animal behaviour, more specifically in relation to our idea of animal-as-scientist. However, we will first look at how rationalism and phenomenology define animal cognition and consciousness.

Contemporary discussion on animal cognition (Griffin, 1978, 1985, 1992; Bekoff, 1998; Allen & Bekoff, 1995; Wynne, 2001), consciousness (for further discussion see Nagel 1974, 1986; Griffin, 1985; Wemelsfelder, 1985; 1997a; 2001; Dawkins, 1993; Chalmers 1995; Dutton and Williams, 2004) and the theory of mind (Premack & Woodruff, 1978) is a departure from an early version of mechanical behaviouristic models of behaviour. Rationalists retain the idea of dualism; so body and mind remain two epistemologically separate entities. Introspection and information processing are currently used to attempt to assess animal cognition and consciousness. These will eventually be addressed along with the development of technology and investigation of the brain. Heyes (1994) defines cognitive behaviour as regulated by representation of its potential outcome. Within this theoretical concept social cognition and the theory of mind are represented by the act of deception. Consciousness, yet another form of introspection, is conditioned by the formation of self-related categories which are manifested through innovative behaviour, tools and language (Meddin, 1979; Peperberg, 1983; Chomsky, 1986). McFarland (1990) defines consciousness as special self awareness that transcends body awareness and involves propositional awareness, which refers back to cognition: "I am the animal aware of the circumstances" (McFarland, 1990). Therefore the proof for someone being conscious is their own self report and an independent verification of their report. The problem, however, is that animals in their natural environment do not use human spoken language, which is the ultimate manifestation of self report. They behave, but since the body is epistemologically separated from the mind, and behaviour mechanisised, scientists within the dualistic tradition enter a loop where the proof seems unattainable. However, in this traditional context, there has been a number of studies on cognition and consciousness in chimpanzees (Meddin, 1979; Premack & Woodruff, 1978), vervet monkeys (Cheney & Seyfarth, 1982), the African Grey parrot (Pepperberg, 1983) and domesticated animals (Wood-Gush et al., 1980; Piggins & Phillips, 1998).

Because of the mind-body split and animals' lack of spoken language and apparent methodological problems with verification, dualists are consistently worried about anthropomorphising animals (Heyes, 1994; Chalmers, 1995). Methodological incapacities within a dualistic theoretical framework can therefore hinder the development of knowledge about animal subjective experience.

On the other hand, phenomenological, relational or interactional approaches (Wemelsfelder 1997a; 2001; Dutton, 2006a) offer a non-traditional definition of consciousness and cognition. There is no body - mind split. The body works as a whole and animal cognition and consciousness are manifested not through internal representations, but through self movement and constant re-evaluation of bodily presence in an interactive environment, either physical or social. Animals' corporeal involvement in interactions is an indicator of their awareness. Intelligence is reflected in the capacity to act and resolve the problems situated outside, in the world, interactively using self directed and expressive behaviours. Gallup (1970, 1977) conducted an experiment on chimpanzees to manifest self recognition using mirror test. He painted red dots on the chimpanzees' forehead and exposed them to mirrors, which the animals used to locate the dots. Gallup concluded that the chimpanzees demonstrated self awareness by their bodily recognition. These claims have been contradicted by an unsuccessful repetition of the same experiment with pigeons (Epstein et al., 1970), and by a theoretical attempt to provide a more mechanical interpretation of the results (Heyes, 1994).

One of the possible reasons for branding animal consciousness as anthropomorphic is to preserve an exclusive "human essence" (Heyes, 1994). The ultimate human exclusivity was introduced in antiquity and reinforced by Descartes' idea of mind, and culminated in contemporary theory of human language (Chomsky, 1986). This theory positions the origins of human language into a species specific mutation (Chomsky, 1986). Therefore other species are categorically excluded from having the capacities for language as a superior form of communication. A number of studies on primates have investigated knowledge and use of human sign or spoken language (Pepperberg, 1986) on other species, predominantly primates (for review see Cavalieri & Singer, 1993; Taylor-Parker & Gibson, 1994).

5.6. Summary and conclusions

In this chapter we first present the idea of animal-as-scientist. We then explain the theoretical background of the model we are proposing and compare it with dualistic perspectives. Finally, we address the question of anthropomorphism. We also tackle the question others have asked but not discussed: are animals personality psychologists (King et al., 1999; Morris et al., 2002)? The model of animal-as-scientist implicitly addresses this question. The model is manifested through animals' interactions with their physical and social environment. Successful interactions imply animals' capacity to assess their environment. As animals do not use verbal language their assessment relies on interactive behaviour, which is as expressive and informative as spoken language (Wittgenstein in Hacker, 1997). The ability to assess others is really the ability to predict or hypothesise the dynamic of their behaviour. In line with the animal-as-scientist model, predictions are based on previous experiences formed into a system of established constructs. When predicting others' behaviour these constructs refer to individuals' consistent behaviour, used in the past. Pervin and John (1997) define personality as behaviour that is consistent over time and situations. An inter-subjectively modified definition of personality would refer to personality traits as behavioural styles that are consistent over time and situations. In intersubjective situations animals need to be able to assess and predict others' behaviour; either for the purpose of survival and obtaining food, as we have seen from the example of lionesses' hunting strategies (Griffin, 1992), or for successful and coherent social dynamic with others, as reported in Sanders' study of dog-human interactions (Sanders, 1992). Therefore, survival and successful social interaction suggest that animals are capable of mutual assessment or construction of consistent behaviours or personality.

The idea of animal-as-scientist is also consistent with most other phenomenological explanations of animal behaviour (Shapiro, 1989, 1990; Sanders 1992; Wemelsfelder, 1997a; Dutton, 2006, in press). It offers a structured explanation of animal individual behaviour both within and between species interactions. It also provides a framework for explaining qualitative and quantitative aspects of behaviour. As explained above, in comparison to dualistic models, the animal-as-scientist model with its phenomenological and intersubjective roots integrates the totality of animal existence, including their emotions, cognition and consciousness.

The issue of anthropomorphism can be addressed from three perspectives within the model. Firstly, knowledge and mutual construing is demonstrated through, meaningful interactions. If our assessment of animals' personality, emotions or awareness were anthropomorphic without any reference to animal experience, then our interactions with animals would be meaningless and unsuccessful. Successful assessment and meaningful interactions depend on familiarity with the individual animals, shared experiences and knowledge about the species specific behaviour. Secondly, the model adopts Kelly's Repertory grid technique (Kelly, 1955) as a method for assessing qualitative aspects of behaviour. Reliability of the assessment is determined by the degree of agreement between the judges. High degree of agreement indicates reliable assessment. From the intersubjective perspective, objectivity is always intersubjective and depends on consensus. Finally, if the assessment of animal personality was just observers' self projection then the observers' personality, empathy or emotional intelligence would affect observers' assessment and degree of agreement. Consequently only observers with similar personality, empathy or emotional intelligence scores would reach an agreement.

In conclusion, shifting the perspective from a dualistic to a relational, phenomenological theoretical tradition provides new ways of reliably applying methods that will assess the qualitative aspects of animal behaviour, including emotions and personality. Nagel (1974) suggests a challenge: in order to study the subjective nature of animals it would be necessary to devise an experimental method that would not rely on imagination or empathy. The method we describe in the next chapter and use in our experimental work, Repertory grid technique, incorporates phenomenological and intersubjective theoretical axioms and offers quantitative analysis of the assessment, and therefore does not rely on imagination and empathy.

CHAPTER 6

Personal Construct Psychology and Repertory Grid Technique

6.1. Introduction

Our initial intention was to apply Kelly's repertory grid technique (RGT) as the method for data collection; however his Personal construct psychology (1955) and the concept of person-as-scientist is also relevant, and has therefore been interwoven in most Chapters. The aim was to use a method similar to free choice profiling (Wemelsfelder et al, 2000), a method that allows the observers to be involved in the assessment of horse personality from the beginning. With RGT the observers had the opportunity to integrate their past experiences and full perception of horse behaviour into personality adjectives that were meaningful to them (Dutton et al., 1997). RGT also allowed us to merge individual lists of adjectives into a uniform scoring list. Items on that scoring list were meaningful to the observers and the data was suitable for application of common statistical tests.

In this chapter we will first look at Kelly's theoretical framework and various applications of the method, then describe RGT and its components, and finally we will discuss modifications we have made to adjust the technique to our experimental design.

6.2. The idea behind Personal Construct Psychology

Kelly developed RGT primarily to use it in clinical settings. The method, also called rep test, is deeply embedded in Kelly's Personal construct theory or Constructive alternativism (Kelly, 1955; Rychlak, 1981). The main premise of his philosophy is not to deny the existence of independent, true reality, but to emphasise the importance of each person's construction of that reality. In order to access that construction we just need to ask the person and they will tell or show us what their reality is. In contrast to personality theories linked to lexical hypothesis and verbal use of language, Kelly does not limit personal construction of reality to language or human beings. Words are only labels for constructs, and some constructs do not have labels. A person structures the meaning of an event or a substance. Even though the structure established by construing can be abstract, that is not necessarily the case. It can also be concrete. For example many constructs are based on behaviour and never communicated using spoken language. Most of those constructs are pre-verbal or nonverbal and deal with individual structures, mostly concerning "physiological construction systems" for example digestion (Kelly, 1963, pp. 51). Therefore animals and infants can have constructs too (Boeree, 2006).

In the previous chapters we have already discussed Kelly's personality theory (PCT) and his idea of person-as-scientist. Clearly Kelly shares a crucial theoretical postulate with other phenomenologists: a person is defined by their anticipation of future events and their active involvement in them (Kelly, 1955). Those anticipations are the consequence of constructive systems. In his analogy with science, forming anticipations or expectations in everyday life is the same as making a hypothesis or predictions in science. The construction system is equivalent to the theory that forms a basis for the hypothesis. These hypotheses, or anticipations or expectations, are constantly tested and verified. In science we do this by designing experiments; in our everyday life by gaining experience from interactions with our social and physical environment. The experience we have gained or experiments we have performed give us the results that we then relate back to our constructive systems or theories, which will be either re-confirmed and accepted, modified and developed, or completely rejected and replaced.

6.3. Kelly's corollaries

Kelly's *fundamental postulate* "A person's processes are psychologically channelized by the ways in which he anticipated events" (Kelly, 1963, pp. 46) expresses our active ability and determination to anticipate future events. Contrary to the early behaviourists, Kelly is convinced that we do not only respond to the environment or are determined by it; but we also anticipate various outcomes and adjust expectations to these variations. However, if the outcome is unusual it will take some time to adjust and change our behaviour and subsequent anticipation.

In order to facilitate his theory, Kelly developed a fundamental postulate and 11 corollaries (Kelly, 1955, 1963). We have already explained some in the previous chapter, however, here is a brief description of all 11 corollaries.

The *construction* corollary refers to our ability to construe anticipations or hypotheses according to our previous experiences. Therefore our construction systems are likely to be conservative and endeavour to replicate what is already known. The essence of the *experience* corollary is our continuous validation of constructs. As a consequence of our experiences the constructs will either be kept unchanged or they will be modified or replaced with new ones if necessary. All constructs are dichotomous or bipolar. The *dichotomy* corollary also emphasises that most of our constructs are non-verbal and that words are only labels; therefore individuals without verbal capacities have them too. This is an indicator that Kelly's constructivism does not involve only cognition, but also emotions, moods and behaviours (Pervin & John, 1997).

The constructs can be organized (*organization* corollary) in an ordered way: they can have either core or peripheral positions. Self-related constructs are usually core as we are mostly reluctant to change them. Peripheral constructs on the other hand are often about the world around us and easily modified or replaced. Our personal organization of ordinate and sub-ordinate constructs will most likely differ from scientific taxonomies or other peoples' organization. We also have constellational or group forming organization of constructs. Therefore the meaning of one construct is automatically associated with several others in the same group. The relationships between those constructs can be tight or loose. Creativity usually helps to loosen the connections and makes it easer for us to change the ties between them. The *range* corollary refers to the limitations of application of some constructs due to our "anticipation of a finite range of events only" (Kelly, 1963, pp 68). In a similar way the *modulation* corollary says that some constructs modulate or are permeable and can be used for various things (e.g. "good-bad"), while others are impermeable and are reserved for very specific things (e.g. most scientific hypotheses are impermeable). Dilation and constriction refer to broadening or narrowing the range of constructs. The most drastic alterations of the range, Kelly suggests, are emotional rather than cognitive. The *choice* corollary again indicates the conservative nature of our construction systems. We choose behaviour that will elaborate our existing construction system, and not change it. Choices about how to construe our reality can therefore be safe or adventurous. Even though we tend to organize our constructs, we can use "a variety of construction subsystems which are inferentially incompatible with each other" (*fragmentation* corollary, Kelly, 1963, pp. 83). As we play different roles in life, we often modify or sometimes inconsistently use our constructs.

The important part of Kelly's theory is a group of corollaries denoting our social interactions. The *individuality* corollary refers to Kelly's postulate that everyone is different: we all have different experiences and consequently different constructions of reality. This is at the same time one of the most basic phenomenological premises.

The concept is similar to Adller's individual traits (Pervin & John, 1997) or Rogers' phenomenal field (1989). The last two are the *commonality* and *sociality* corollaries. The both refer to the sharing of our constructs with others. By introducing these two corollaries Kelly departs from the orthodox phenomenology developed by Husserl (Crossley, 1996). The *commonality* corollary states that the degree of closeness within a community, culture or a sub-culture is reflected through similarities in our personal construct systems. Validation from others is something that is important and for which we are constantly striving. The *sociality* corollary also includes the concept of empathy or emotional intelligence. Kelly says that even if we do not really share that much with another person, we can still put ourselves in their shoes and get a feeling of how they feel and so construe their reality.

6.4. Applications of repertory grid technique

Repertory grid technique (RGT) was developed to formalize the various stages of construing. It has been used as a tool for investigating individuals' constructs, their relations and flexibility (Kelly, 1955; Cooper, 1998). It has mostly been utilised in clinical psychology (Landfield & Epting, 1987). An example of a semi-clinical setting is seen in a study by Smail (1972), who used RGT to measure empathy in a therapeutic group. The main emphasis was on understanding patients from their point of view and therefore adjusting the methods to the subjects. The need to understand other's points of view and how they see the world around them has not been limited to psychology only. The most frequent application outside the clinical setting has been in food science (Schutz, 1988). Consumers' choice of different foods (Schutz, 1988) and drinks e.g. lagers (Gains & Thompson, 1990) is based on their qualitative characteristics, contextual factors and the consumers' habits, culture, personality, mood and physiology (Gains, 1994). Gains points out that in market research many approaches (pre-structured questionnaires, focus groups and free elicitation) suffer from a poor response rate, interviewer bias or difficulties describing product characteristics especially if they are assessed separately, and that RGT avoids these problems. It also reflects the full range of differences between the objects of study (Gains, 1994)

For similar reasons Dutton et al. (1997) used RGT for the first time in animal behaviour studies to assess the personality of chimpanzees. Their intention was to apply a method that would reflect observers' individual construction of chimpanzee personality. Prior to that, adjectives were usually borrowed from human personality studies or generated by species experts or the experimenter, not by the people who rated the animals. Dutton, however, aimed to produce chimpanzees' personality dimensions consisting of traits based on the impressions of experienced observers who knew the individual animals (Dutton et al., 1997). As she points out these descriptors were based on observers' knowledge of the species, "subjective impressions" and memory of actual behavioural acts of the chimpanzees. She argues that individually elicited constructs provided a more accurate assessment of the chimpanzees' personality which may help in predicting their future behaviour and concludes that the method is appropriate for quantification of intuitive personality evaluation. The only other application of RGT to animal behaviour was done with pigs (Grajfoner et al, 2002). The method was modified, using pairs of elements (pigs) for elicitation. The use of dyads (pairs) instead of triads of elements was suggested where triadic elicitation would be too confusing for observers (Ryle and Lunghi, 1970). This will be discussed further in the next section. The method was found to produce significant inter-rater reliability. Its validity was examined by cross-validation across two experiments. The same animals had previously been rated by a different group of observers, using a different method, free choice profiling (Wemelsfelder et al, 2000). Results from the two studies were significantly correlated.

Repertory grid technique has therefore been applied in many areas. Fransella and Bannister (1977) and Fransella et al. (2004), however, emphasise two important factors when choosing the items or elements being rated: first, they should be relevant to participants and second, they should also be representative of the pool of elements. The participants should be able to relate their self-constructed role to the items or elements in order to assess them.

6.5. The constituents and stages of the rep test and their modification in our study

The main constituents of RGT or of the rep test are elements, subjects and constructs. In our study the elements were horses, the subjects were observers and the constructs were the personality adjectives observers generated for the horses, either by watching them on video or interacting with them over a longer period of time. We shall use this terminology interchangeably with observers or raters for subjects and horses or animals for elements. We will also use both RGT and rep test for repertory grid technique.

The rep test consists of three stages: the first stage is elicitation of bipolar constructs; the second is elaboration of the constructs to increase understanding and clarification; the third stage is rating of elements on the constructs the subject has provided. During the second stage, identical constructs are usually eliminated. We will now describe the three stages and discuss the modifications we have made in our application of the method.

Construct elicitation is usually done in triads. The subjects are asked to compare two elements and contrast them with a third. During the comparison one pole of the construct emerges. By adding the third element and contrasting it to the first two, the subject provides the opposite pole of the construct. The names of the elements, which can be people, foods or animals (for examples see Fransella & Bannister, 2004) are usually written on individual cards. There are various ways of presenting the triads (see Fransella & Bannister, 2004). The ones interesting for us were the minimum context form, the full contecst form and the sequential form. The minimum context form of elicitation requires the subject to generate the constructs using three elements to compare and contrast at the same time. When the subject has exhausted the elicitation of constructs by comparing and contrasting, the first triad is replaced with the second triad and so on. The full context form on the other hand postulates that all cards, not just triads, are presented to the subject at the same time. The subject then needs to group the cards and generate constructs by comparing and contrasting these groups.

The form we used in our study was the sequential form of elicitation. This form is similar to the minimum context form. The only difference is that not all three cards or triads are replaced by new ones, but just one new card or element is systematically added to the existing triad as one of the old ones is removed. Our sequential form was modified so that we used dyads or pairs instead of triads of elements or horses. Sometimes comparing three elements is too demanding, therefore using dyads of elements is an acceptable option. This has been done before with children (Allison, 1972 in Fransella & Bannister, 1977), when exploring interpersonal relationships (Ryle & Lunghi, 1970) or rating acquaintances (Fjeld & Landfield, 1961). Our choice of using dyads turned out to be a good decision especially with video observations, where the observers did not know the horses and already struggled to compare and contrast two elements.

6.6. Eliciting or providing the constructs.

Manuals for using rep grid are fairly flexible regarding the provision or elicitation of the constructs. Fransella and Bannister (1977) suggest that the provision of bipolar descriptors can be useful as sometimes the subjects cannot verbalize their own constructs. However, we need to make sure that the labels we provide are meaningful to the subject. Provided constructs can vary from those that would be used by the subjects anyway to those that are completely unacceptable. We need to interview the subject to find out which of the provided constructs are acceptable and which are meaningless to them. Our second stage of the rep test were interviews in which the observers elaborated their constructs and identified duplicates.

Fransella et al. (2004) indicate that it is not uncommon to use provided constructs. The best way is to collect the constructs most commonly used by the experimental or social group to which the subject belongs. Fransella and Bannister (1977) argue that it is safe to assume that constructs collected in that way will be meaningful to the individual (Fransella & Bannister, 1977, p. 19). They also say that constructs obtained in that way are not just elicited or provided. In our study each subject generated their own list of constructs. These were then merged by the experimenter into a uniform scoring list. The constructs on that uniform scoring list were those elicited by most subjects.

Personal construct psychology argues that a person will regard their own constructs as more important (Adams-Webber, 1970; Fransella & Bannister, 1977). It has also been suggested that elicited constructs will provide more reliable results (Fransella et al., 2004). This premise has been echoed in the integrative assessment of animal behaviour (Wemelsfelder et al., 2000). In the experimental part of our project we investigated whether individually elicited constructs are indeed more reliable. Our uniform scoring lists were composed by merging individual sets of constructs. Each uniform scoring list that we composed consisted of both constructs that were elicited by all observers and constructs elicited by just a few. Therefore we could explore the differences in the degree of agreement between the observers on elicited and "provided" constructs.

6.7. Summary and conclusions

Kelly developed RGT to understand how others construe their social and physical environment. In our study we aimed to understand how observers construed horses and differences in their behaviour. Repertory grid technique is a method that allows an insight not only into horse personalities and observers' constructions of these, but also into the interactions between the two. The method can be and has been applied without including the Personal construct theory. That has been done in a number of predominantly food related studies, as reviewed in this chapter. Kelly's theory, however, is important for this project, so it has been integrated into various parts of the thesis.

In conclusion, this chapter has introduced the basic premises of Kelly's Personal construct theory and RGT. These premises or corollaries address different aspects of construing. The method promotes idiographic principles with individually elicited constructs, and at the same time allows a degree of standardization (Adams-Weber, 1970). We have described variations in the elicitation stage and explained the modifications we made for our project. Instead of triads we used dyads of elements with a sequential form. The use of RGT in assessing animal personality (Dutton et al., 1997) and behavioural expressions (Grajfoner et al., 2002) has produced consistent results that are reliable and valid. A beneficial outcome of using the method is a pool of personality adjectives directly relevant to horses which can be used for future research.

CHAPTER 7

Experimental Studies

Overall structure of the experiments

This project includes two studies: the first or the pilot study and the second or the main study. The first study was conducted at two stables with 21 horses. Forty-five participants were observing and rated the horses, eight participants knew the horses and 36 participants were not familiar with them (Figure 7.1).

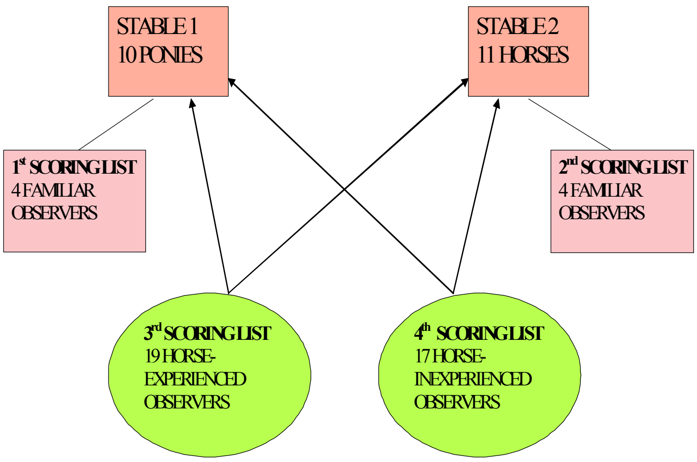


Figure 7.1

Horses and observers in Study 1

In the second study there were three stables with 38 horses. Thirty-four observers participated in the study. Twenty-four were familiar with the horses and ten did not know the horses at all (Figure 7.2).

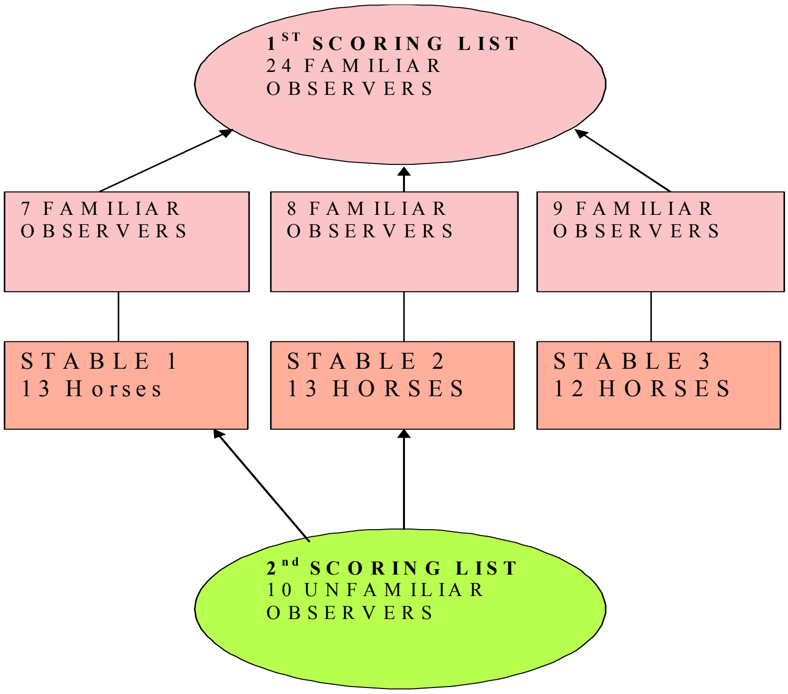


Figure 7.2

Horses and observers in Study 2

Familiar observers rated the horses based on their previous experience with the horses. Inexperienced observers watched videos of individual horses interacting with a stranger. Video observation in the second study also included a novel object test with an open umbrella. Results from Study 2 are presented in Chapter 8.

Observers’ personality, empathy and emotional intelligence were an integral part of the experiments. We were interested in how and to what extent these factors determined their rating. Results from this part of the project are presented in Chapter 9.

Horse Personality – Study 1

7.1. Methods

7.1.1. Animals and housing

For the first study 21 horses from two riding schools were used. One was a riding school for the disabled (Stable 1), the other was a regular riding school (Stable 2). The horses from Stable 1 were different types of pony geldings, aged eight to 24 years and had been used in Stable 1 from five months to 17 years (Table 7.1). The horses were usually kept in separate pens during the night and used for inside and outside riding during the day.

Table 7.1

Horses - Stable 1

HORSE	AGE–years (2001)	SEX	BREED
Eddie	22	Gelding	Connemara x
Ginger	24	Gelding	Welsh x
Zero	16	Gelding	Highland
Hamish	18	Gelding	Connemara x
Sam	12	Gelding	x Breed
Stroller	11	Gelding	Cob
Bobby	9	Gelding	Irish Draught
Sandie	8	Gelding	Cob x
Fudge	8	Gelding	Dalos x

Cuillin	7	Gelding	Highland

The horses at Stable 2 were eight geldings and three mares. They were thoroughbred and thoroughbred crosses, aged six to 19 years (Table 7.2). Three horses were not used for public riding. All the horses were kept in separate pens and used for riding, mainly inside.

Table 7.2
Horses - Stable 2

HORSE	AGE–years (2001)	SEX	BREED
Dolce	6.5	Gelding	Thoroughbred
Rob	9	Gelding	Thoroughbred
Slotty	17	Gelding	Thoroughbred
Stroppy	6	Gelding	Thoroughbred x Irish
Lyric	7	Mare	Thoroughbred x
Reiver	13	Gelding	Highland
Molly	16	Mare	Thoroughbred x
Misty	19	Gelding	Connemara
Appley	11	Gelding	Connemara
Lizzy	7	Mare	Thoroughbred x Warmblood
Coco	17	Gelding	Irish x

7.1.2. The observers

In this study we used three groups of participants: a) "familiar" observers were familiar with the horses, and two groups of unfamiliar observers: b) "experienced" observers who had had previous experience with horses but did not know the horses in the study, and c) "inexperienced" had had no experience with horses at all. Because of the importance of understanding and using language appropriately for describing horses' behaviour, the participants were native English speakers. As all familiar observers were female we recruited only women for unfamiliar groups to minimise inter-group variations due to gender.

In the familiar group there were five observers who initially started the experiment at each riding school. In Stable 1 one observer was excluded from statistical analysis due to her short attention span during the rating and potentially unreliable results; in Stable 2 one observer was not included in the analysis due to only partial familiarity with some horses. Therefore eight familiar observers completed the study. The participants were all women aged 16 to 56. They worked as trainers or grooms. They had observed and interacted with the horses in a variety of situations and they had known the animals from 5 months to 17 years.

There were two groups of unfamiliar observers in the video part of the experiment. In the group of experienced observers there were 19 females; six were veterinary students, 13 psychology students. They all had previous experience with horses but were not familiar with the animals they observed and rated. The second group of observers were 17 females recruited from the volunteer panel from the Psychology Department at Edinburgh University, aged 25-55, and labelled as "inexperienced". The condition for participation was that they had had no previous experience with horses at all.

Collecting information about the observers

The observers were asked to complete the NEO FFI personality test (Costa and McCrae, 1992) and the observer information form, providing some basic personal information: sex, age, level of education and whether they were vegetarians or not. During the elicitation part of the experiment they were asked to complete the Animal Empathy Test (Paul, 2000). Before rating the horses they were also asked to complete the UWIST mood test (Matthews et al., 1990).

7.1.3. Experimental procedure

In the first part of the study the horses were rated by the familiar observers. In the second part the horses were recorded on video, interacting with a human interactor in a human-interaction test (see Wemelsfelder et al., 2000). These videos were the source of rating for inexperienced observers.

7.1.3.1. Repertory Grid Technique

The repertory grid technique is a method used mostly for collecting qualitative data. It consists of two parts: elicitation of bipolar constructs (e.g. friendly-unfriendly) and rating: observers rate each individual horse on these constructs (see Chapter 6 for general description). The method we used for familiar and unfamiliar observers was slightly different, therefore we describe each procedure separately.

7.1.3.1.1. Rating by the familiar observers

7.1.3.1.1.1. The elicitation procedure

The purpose of the elicitation procedure was to generate personality descriptors which were meaningful to the participants when describing the behaviour of the horses. The names of the horses were printed on separate cards. A combination of the Minimum Context Card Form and the Sequential Form was used (see section 6.5. for further explanation). Participants were presented with the names of two horses and asked to describe the differences and similarities between them. The first card was replaced with a new one and the observers had to compare and contrast the new pair. The card with the name of a horse that was paired with two other horses was replaced by a new card and so on. The elicitation was done individually and it lasted from 1.5 to 2 hours per person. The experimenter did not know the horses at all and chose the cards randomly. In the process of contrasting and comparing the observers were asked to describe similarities and differences between the two horses in terms of their overall behaviour, consistent style of interacting with the environment and personality traits. They also had to provide an opposite term for each descriptor they elicited.

7.1.3.1.1.2. Interviewing methods

After completing the elicitation procedure, each observer was asked to define the adjectives or to rephrase them. The purpose of that part was to elaborate what the observers meant by the constructs. According to Kelly (1955) words are only labels for underlying constructs, so it was important to synchronise the individual elicitation lists for the next stage of the experiment. The participants provided synonyms or a short explanation for their constructs (for an individual elicitation form with elaborations see Appendix 1). Following that the participants were asked to identify identical constructs and finally to choose the 15 most important constructs they use to describe the horses.

7.1.3.1.1.3. Composing the uniform list of constructs

After the elicitation and the interviews, we composed a uniform list of bipolar constructs, based on the individual elicitations. There were two familiar groups of observers, so we composed two separate uniform lists, one per group. The constructs were ordered according to how many participants used them. For example, the constructs produced by everybody were put at the top of the list. Descriptors listed by one participant but mentioned by others were put at the end. Highly idiosyncratic constructs were eliminated. The uniform list of constructs for Stable 1 had 32 items (Table 7.3), and the list for Stable 2 consisted of 28 items (Table 7.4).

Table 7.3

List of constructs from Stable 1 by familiar observers

1. CONFIDENT – UNSURE
2. HAPPY – UNHAPPY
3. LAID-BACK – WORRIER
4. RELIABLE – UNRELIABLE
5. BULLY – BULLIED
6. INTELLIGENT – STUPID
7. WILLING TO WORK – UNWILLING TO WORK
8. STEADY – FAST
9. BOLD – TIMID
10. FRIENDLY – UNFRIENDLY
11. FORWARD GOING – SLOW
12. DOES NOT LIKE SMALL AREAS – NOT WORRIED ABOUT IT
13. PUSHY – DOCILE
14. TRUSTFUL – SUSPICIOUS
15. AFFECTIONATE – ALOOF
16. BOSSY – LIKES TO FOLLOW
17. OBEDIENT – DISOBEDIENT
18. EASY TO LEAD – STUBBORN
19. GOOD LEAD HORSE – PREFERS FOLLOWING
20. INQUISITIVE – NOT INQUISITIVE
21. RESPONSIVE – UNRESPONSIVE
22. ACCEPTING – NOT ACCEPTING
23. SOCIABLE – LONER
24. LIVELY – SLUGGISH
25. PLAYFUL – DULL
26. QUIET – EXCITABLE
27. PATIENT – IMPATIENT
28. GOOD NATURED – STROPPY
29. RELAXED – UPTIGHT
30. ALERT – APATHETIC
31. SENSITIVE – UNFEELING
32. GREEDY – NOT GREEDY

During the interview the participants were asked to identify the most useful constructs for describing horses' personality traits. Constructs in bold were identified as the most useful by at least three participants (Table 7.3).

Table 7.4

List of constructs from Stable 2 by familiar observers

1. BOLD – SHY
2. CLEVER – STUPID
3. SOCIABLE - ANTI-SOCIAL
4. CHEEKY – TIMID
5. FRIENDLY – UNFRIENDLY
6. CALM – EXCITABLE
7. LAID-BACK – HIGHLY STRUNG
8. AFFECTIONATE – ALOOF
9. CONFIDENT – NERVOUS
10. WILLING TO PLEASE – STUBBORN
11. SAFE – SPOOKY
12. CONTENT – UNSETTLED
13. OBEDIENT – DISOBEDIENT
14. INTERESTED – DISINTERESTED
15. LIVELY – DULL
16. EASY TO WORK WITH – DIFFICULT
17. PLAYFUL – BORING
18. SECURE – INSECURE
19. PREDICTABLE – UNPREDICTABLE
20. HAPPY – UNHAPPY
21. WELL BEHAVED – NAUGHTY
22. SENSITIVE – UNRESPONSIVE
23. FRESH – LAZY
24. WELL MANNERED – IGNORANT
25. QUIET – IMPATIENT
26. GENTLE – AGGRESSIVE
27. RELIABLE – UNTRUSTWORTHY
28. GOOD NATURED – STROPPY

Legend

Constructs in bold were defined as most useful by the observers (Table 7.4).

7.1.3.1.1.4. Scoring

The scoring was the third and the last part of the RGT. It took about an hour and it was done in a group at each stable. The uniform scoring list consisted of 32 bipolar constructs for Stable 1 (Appendix 2) and 28 constructs for Stable 2 (Appendix 3). The participants rated the horses on a 5-point scale per construct (e.g. FRIENDLY 1 2 3 4 5 UNFRIENDLY). They were instructed to encircle an appropriate number on the scale. If they thought that a horse was very friendly they encircled 1, if they thought it was moderately friendly they encircled 2, if they thought it was neither friendly nor unfriendly they encircled 3 and so on. They were instructed to think about the horses' behaviour across various situations, and rate them accordingly. It was emphasised that they should think about the horse's interactions with other horses and with people. At the end of the scoring the participants were asked to identify the 15 most useful constructs for describing individual differences in horses. They were also invited to identify the constructs which they thought were not applicable to describe horses. All the constructs were kept and used for the analysis. The raters did not collaborate or communicate with other raters prior or during scoring.

7.1.3.1.2. Unfamiliar observers - Video part

The experimental procedure for this part was the same for both stables. The procedure involved recording a video, which was edited and shown to unfamiliar observers. The recording of the video took place in one day, with a duration of two and a half hours per stable. The horses were not previously trained for the experiment. The video at Stable 1 was made in an outside paddock and the video at Stable 2 in an inside arena. The horses were not usually kept in those areas, therefore the locations were fairly new to the horses. To create an experimental area, both paddock and the arena were partially restricted by a rope and posts to approximately 10m x 5m. A human interactor, who was not familiar with the horses, was standing in the middle of the restricted area. Horses were brought individually into the area by a familiar person and released as soon as they entered the area. The area was then closed and the familiar person left the site. During the recording of the video no other people or horses were on the site. The interaction was not standardised, although there were some basic rules (see Wemelsfelder et al., 2000): the human interactor waited until the horse approached her and initiated the interaction. If the horse looked at the interactor she would extend her hand and wait for the horse to approach her. If the horse started the interaction, she would touch their forehead or neck and if the horse showed more interest in interaction, she would use both hands to pat their head and neck. If the horse stopped showing interest the interactor stopped patting the horse and only extended her hand if the horse looked at her.

The video clips were edited and separated into two parts. One minute inserts were used for elicitation, and three minute videos with two minutes blank screen for scoring.

7.1.3.1.2.1. The elicitation procedure

The cards with horses' names were replaced by one-minute videos of each individual horse interacting with a human interactor. We applied the sequential form with a dyadic elicitation: observers were presented with the videos of two horses, one following the other with a 5 second black screen in between. The most contrasting horses were paired up according to the experimenter's decision. The participants first watched the videos of horses from Stable 1 and after a short break they watched the videos of horses from Stable 2. The observers were instructed to describe the horse personality on the basis of their style of interaction with the human interactor and describe similarities or differences between the two horses. After that the second pair of horses was presented to the observers and so on. The elicitation lasted up to 90 minutes per observer and they were all done individually. The observers watched the videos only once.

7.1.3.1.2.2. Interviewing methods

The interviews were done individually and they took up to an hour. Immediately after the elicitation the observers were asked to elaborate the constructs by providing a short definition, explanation or synonym. They were also asked to point out identical constructs and to highlight the 10 to 15 constructs they thought were the most useful for describing the horses (An example of an elicitation form with elaborations is included in Appendix 4).

7.1.3.1.2.3. Composing the uniform scoring lists

The two unfamiliar groups were treated as two separate observer groups and therefore had their own scoring list each.

7.1.3.1.2.3.1. Experienced observers

A uniform list of constructs was composed by the experimenter based on individually elicited constructs. The scoring list used to assess the personality of horses at Stable 1 and Stable 2 comprised 31 items or pairs of adjectives (Table 7.5). These constructs were put in order according to the number of participants who generated them. For example, constructs common to everybody ("interested-disinterested") were put at the beginning of the scoring list. After that constructs generated by at least 1/3 of the experienced observers were included. Constructs generated only by one observer and mentioned during the elicitation or interviews also by others were included at the end of the scoring list. Highly idiomatic constructs or repetitions were eliminated. Only the first pair was generated by all the participants. The first 16 constructs were generated by at least one third (7) of the observers, and the following 10 by at least half of the participants (10). The last five constructs were elicited by one or two observers.

Table 7.5

List of constructs generated by experienced observers (19) for both groups of horses

1. INTERESTED	DISINTERESTED
2. BOLD	TIMID
3. FRIENDLY	UNFRIENDLY
4. CONFIDENT	NERVOUS
5. BORED	ALERT
6. RELAXED	TENSE
7. LAID-BACK	SPIRITED
8. CALM	EXCITED
9. UNCONCERNED	WORRIED
10. CONTENT	UNHAPPY
11. SOCIABLE	UNSOCIABLE
12. ACTIVE	PASSIVE
13. SECURE	FRIGHTENED
14. SPOOKY	TRUSTWORTHY
15. SURE	UNSURE
16. QUIET	ENERGETIC
17. TRUSTING	UNTRUSTING
18. PLACID	HOT TEMPERED
19. SAFE	JUMPY
20. DOCILE	HIGHLY STRUNG
21. SETTLED	UNSETTLED
22. PLAYFUL	SERIOUS
23. AWARE	OBLIVIOUS
24. GENTLE	AGGRESSIVE
25. AFFECTIONATE	UNAFFECTIONATE
26. STEADY	FLIGHTY
27. GOOD NATURED	DIFFICULT
28. INTELLIGENT	STUPID
29. EASY GOING	STUBBORN
30. EXPERIENCED	INEXPERIENCED
31. PATIENT	IMPATIENT

7.1.3.1.2.3.2. Inexperienced observers

The list generated by the unfamiliar observers consisted of 26 items or pairs of adjectives (Table 7.6). The constructs were put in order according to how many participants elicited them. For example, a construct generated by all inexperienced participants ("interested-disinterested") was put at the beginning of the scoring list. Constructs generated by one or two observers were put at the end of the list. As with previous observer groups, highly idiosyncratic constructs were not included in the list. In total there were first the 13 constructs generated by at least one third (six) of the participants, the first seven having been generated by at least half (nine) of the participants.

Table 7.6

List of constructs generated by inexperienced observers (17) for both groups of horses

1.	INTERESTED	DISINTERESTED
2.	FRIENDLY	UNFRIENDLY
3.	CALM	AGITATED
4.	CONFIDENT	NERVOUS
5.	HAPPY	UNHAPPY
6.	PLACID	HIGHLY STRUNG
7.	PLAYFUL	SERIOUS
8.	ACTIVE	PASSIVE
9.	SURE	UNSURE
10.	BORED	INTERESTED
11.	DOCILE	SPIRITED
12.	ALERT	TIRED
13.	INTERACTIVE	DISTANT
14.	TOLERANT	INTOLERANT
15.	RELAXED	TENSE
16.	SECURE	INSECURE
17.	TRUSTING	DISTRUSTFUL
18.	EXTROVERTED	INTROVERTED
19.	APPROACHABLE	STAND-OFFISH

20.	AFFECTIONATE	UNAFFECTIONATE
21.	BOLD	TIMID
22.	ACCEPTING	REJECTING
23.	SOCIABLE	UNSOCIABLE
24.	PATIENT	IMPATIENT
25.	GENTLE	HARSH
26.	WELL BEHAVED	STUBBORN

Legend

The first 13 constructs were generated by six or more participants, and the first seven by nine or more participants. The last four constructs were generated by two participants.

7.1.3.1.2.4. Scoring

The scoring procedure took about one hour. There were 31 bipolar constructs on the experienced scoring list and 26 constructs on the inexperienced scoring list (These scoring lists are included in Appendix 5 and 6). The participants rated the horses on a 5-point scale, having one pole of the constructs on one side and the second pole on the other side (e.g. FRIENDLY 1 2 3 4 5 UNFRIENDLY). They were instructed to circle the appropriate number on the scale. If they thought that a horse was very friendly they had to circle 1, if they thought it was moderately friendly they circled 2, if they thought it was neither friendly nor unfriendly they circled 3 and so on. At the end of the scoring the participants were asked to identify the 10-15 most useful constructs for describing individual differences in the horses. They were also invited to identify constructs which they think were not applicable to describe the horses. All the constructs were included in the analysis.

7.1.4. Data Analysis

The Kendall's Coefficient of Concordance tests inter-observer agreement. We applied the same test when testing observers' agreement when the horses were rated on elicited or partially provided constructs and on the most useful constructs and the rest. We used the mean scores with standard deviation for graphical presentation of personality trait profiles in each horse. Observers' consistency in applying individual constructs was also tested using Kendall's W test of concordance. ANOVA, and t-tests were used to identify the differences in the mean Kendall's W scores between any of the observer groups. Multidimensional (MDS) scaling was applied to discover the main horse personality dimensions. Linear regression helped us to identify them. To determine if there were any significant correlations between the dimensions produced by the three observer groups we calculated Spearman correlation coefficients for MDS scores. SPSS 12.0, SigmaPlot 7 and Excel 2003 statistical packages were used to analyse the data and to produce the figures and tables.

7.2. Results

7.2.1. Inter-observer agreement

First we calculated the degree of agreement within each observer group: familiar, experienced and inexperienced. Secondly, we looked at the difference in the degree of agreement when the horses were rated on elicited and partially provided constructs, and on the most useful constructs and the rest. Thirdly, we calculated Kendall's W to determine observers' consistency in using the constructs. Finally, we calculated the mean constructs' scores with standard deviations for each horse and presented horse personality profiles.

7.2.1.1. Observers' agreement on all constructs

7.2.1.1.1. Stable 1

We calculated Kendall's coefficient of concordance W to determine overall agreement within each group of observers when scoring the horses on all elicited constructs. When rating horses from the first stable there was significant agreement within all groups of observers for all but one horse (Figures 7.3 and 7.4). The exact Kendall's W values are included in Appendix 7.

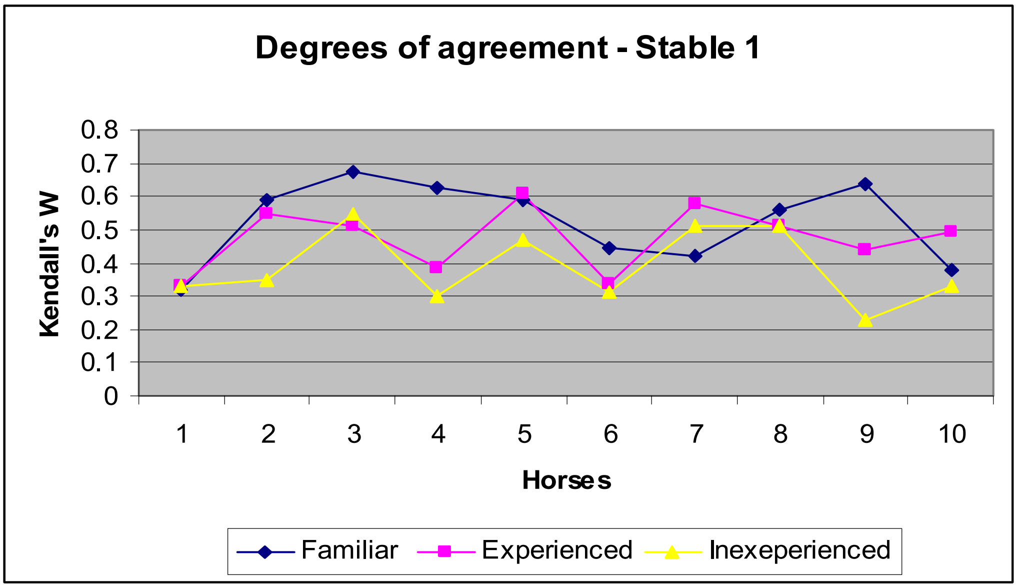


Figure 7.3

Observers scoring horses on all constructs – Stable 1

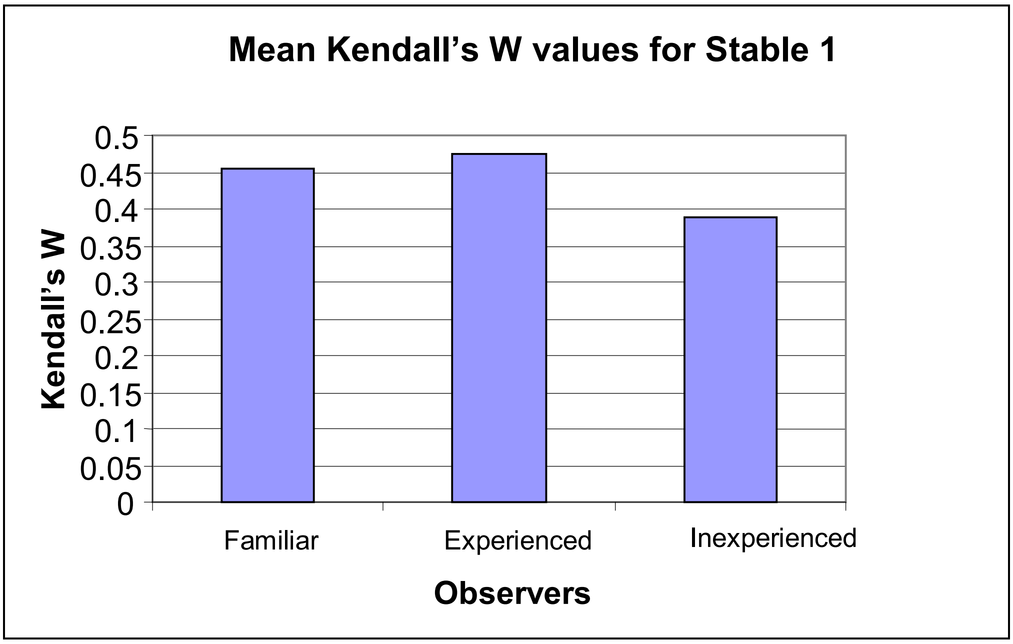


Figure 7.4

Mean Kendall's W values for three observer groups – Stable 1

To establish if there was a difference in Kendall's W between the three observer groups, we did ANOVA, which showed significant difference in the degree of agreement between the three groups F(7,1) 5.79, $p<0.05$. Further, paired t-tests demonstrated no significant difference between familiar observers and experienced observers ($t=1.19$, $p>0.05$). However there was significant difference in the degree of agreement between familiar and inexperienced ($t=2.78$, $p<0.05$) and experienced and inexperienced ($t=3.1$, $p<0.05$) observers.

7.2.1.1.2. Stable 2

Kendall's coefficient of concordance W was also calculated to determine whether there was an overall agreement within all three groups of observers when rating horses from the second stable. The degree of agreement was highly significant for the majority of horses (Figures 7.5 and 7.6). The exact Kendall's W values are included in Appendix 8.

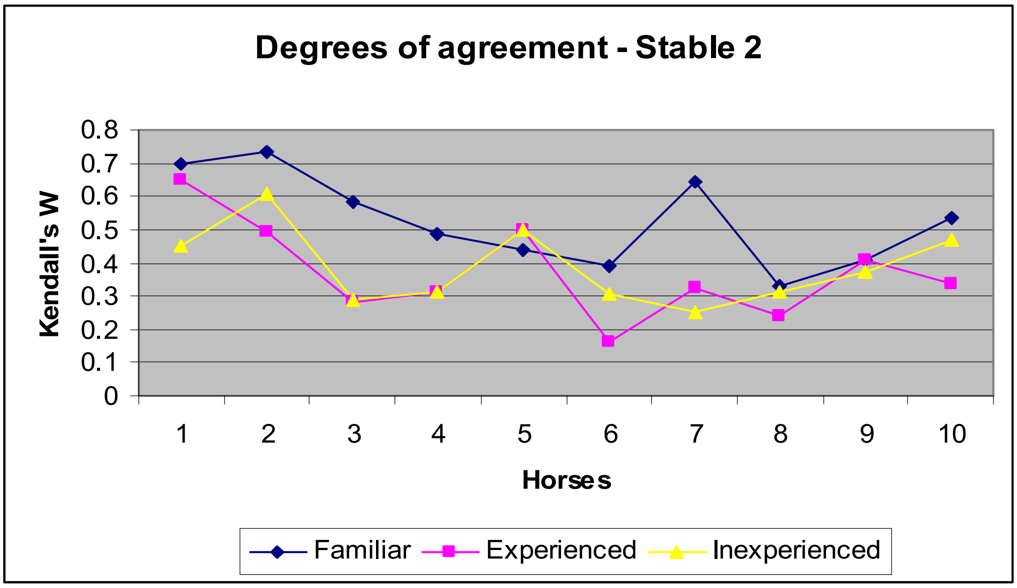


Figure 7.5

Observers scoring horses on all constructs – Stable 2

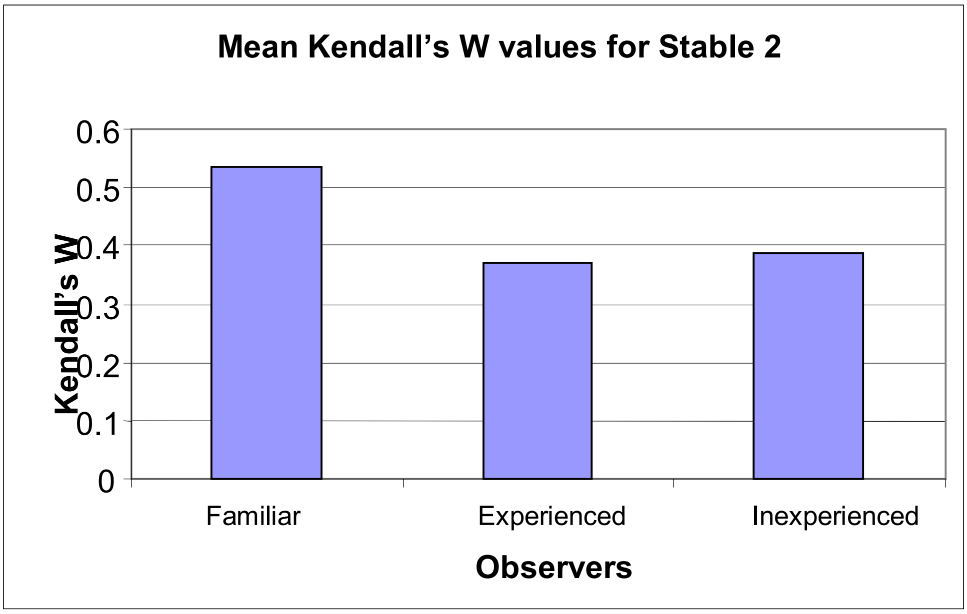


Figure 7.6
Mean Kendall's W values for three observer groups – Stable 2

ANOVA showed significant difference in the mean values of Kendall's W between the three observer groups ($F(7,1)$ 9.16, $p<0.01$). Paired t-tests showed there were significant differences between familiar and experienced observers ($t=3.79$, $p<0.01$) and familiar and inexperienced ($t=3.14$, $p<0.05$), but not between experienced and inexperienced ($t=-0.47$, $p>0.05$) observers.

7.2.1.2. Difference between elicited and partially provided constructs

We were interested to find out if the observers agreed more on the constructs they generated themselves. So we first split the scoring list in half. The constructs were divided according to how many observers generated them. In the first half there were constructs elicited by three quarters or more (familiar) or half or more (unfamiliar) of the observers. In the second part there were constructs generated by less than three quarters of familiar or half of all the unfamiliar observers respectively. We calculated Kendall's W separately for the first and the second halves of the constructs on the scoring lists.

7.2.1.2.1. Stable 1

Kendall's W showed a tendency to higher agreement when the horses were scored on the constructs provided by all the participants in the familiar ($t=2.24$, $p=0.05$) and inexperienced groups ($t=4.89$, $p<0.001$) but not in the experienced group ($t=0.7$, $p>0.05$) (Figure 7.7).

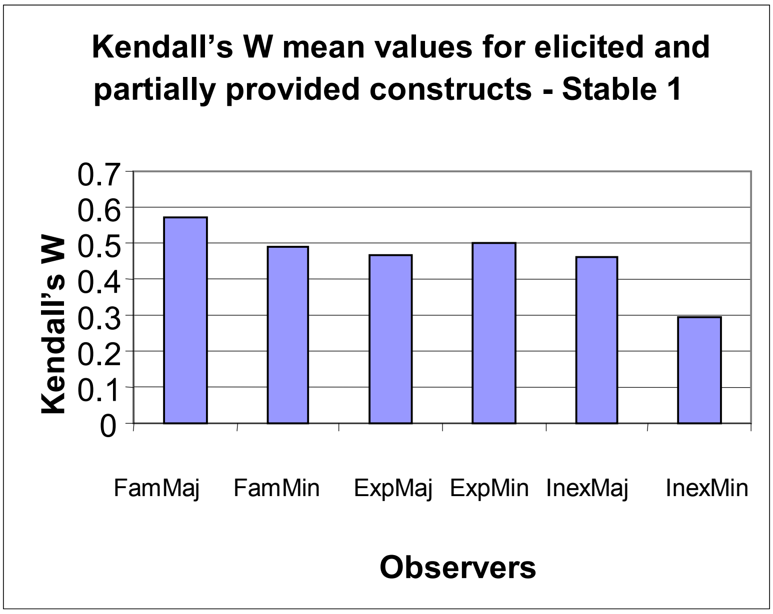


Figure 7.7

Kendall's W mean values for elicited and partially provided constructs Stable 1

Legend: FamMaj (Familiar observers majority), FamMin (Familiar observers minority), ExpMaj (Experienced observers majority), ExpMin (Experienced observers minority), InexMaj (Inexperienced observers majority), InexMin (Inexperienced observers minority).

7.2.1.2.2. Stable 2

In contrast to Stable 1, for Stable 2 only inexperienced observers had a significantly higher degree of agreement when rating horses on generated constructs (t=5.2, p<0.001). There was no difference between familiar and experienced observers (Table 7.8). The exact Kendall's W values for both stables are included in Appendix 9.

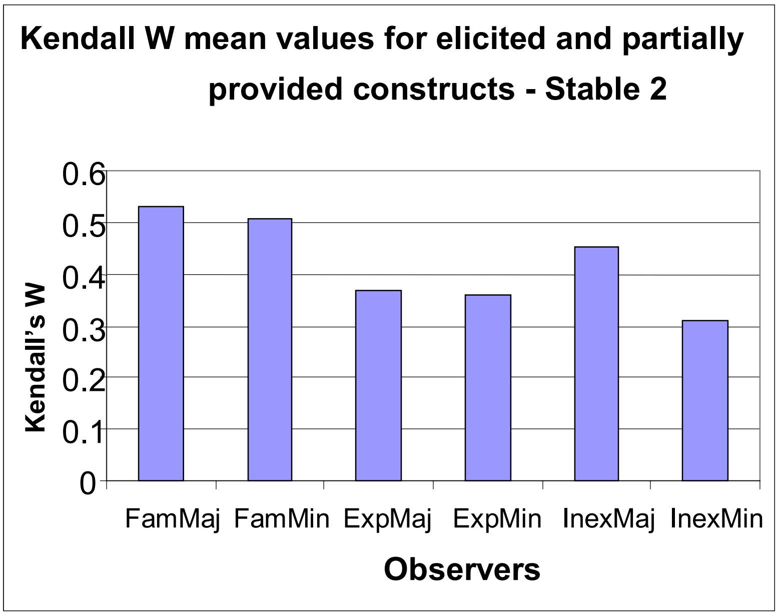


Figure 7.8

Kendall's W mean values for elicited and partially provided constructs Stable 1 and Stable 2

Legend: FamMaj (Familiar observers majority), FamMin (Familiar observers minority), ExpMaj (Experienced observers majority), ExpMin (Experienced observers minority), InexMaj (Inexperienced observers majority), InexMin (Inexperienced observers minority).

7.2.1.3. Difference between the most useful constructs and the rest

The observers were asked to identify the constructs they find the most useful for describing horse personality and horses' styles of interacting. We were expecting a higher degree of agreement on constructs defined as most useful. Kendall's W, however, shows that familiar and experienced observers had a slightly lower degree of agreement on those descriptors in both stables (Figures 7.9 and 7.10).

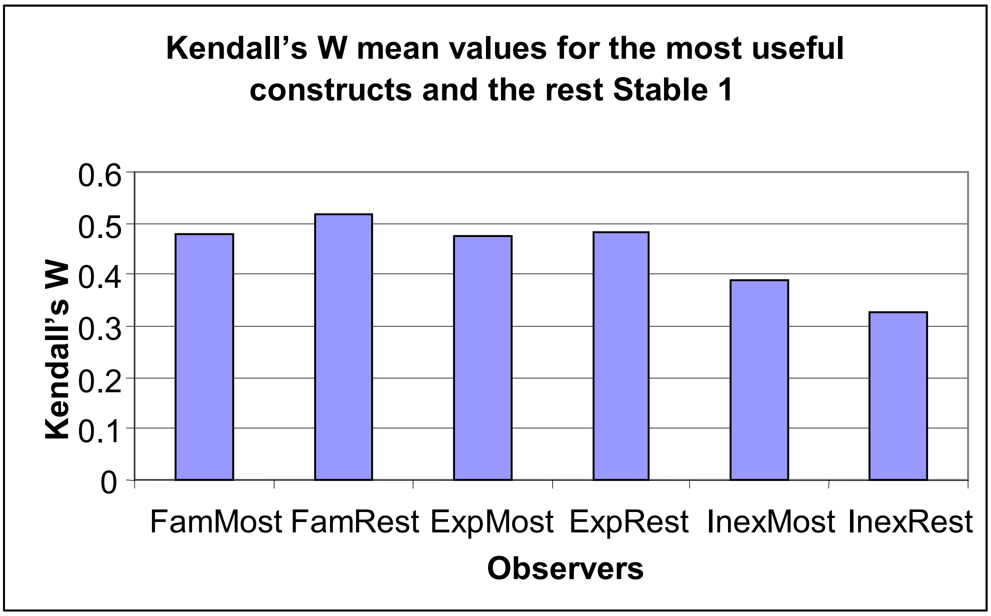


Figure 7.9
Kendall's W values for the most useful constructs and the rest for Stable 1

Inexperienced observers on the other hand had a slightly higher degree of agreement on the most useful constructs (Exact Kendall's W values are included in Appendix 10).

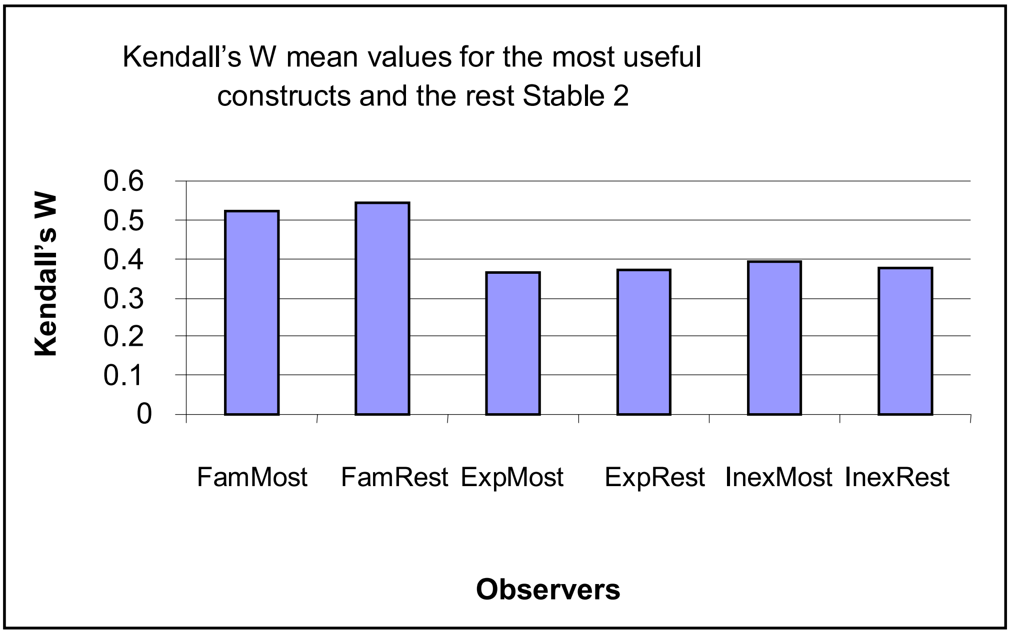


Figure 7.10

Kendall's W values for the most useful constructs and the rest for Stable 2

There was no pattern in the differences in the degree of agreement between the most useful constructs and the rest. Paired T-tests showed no significant difference between the two set of constructs ($p>0.05$) in any observer group.

7.2.1.4. Consistency of scoring horses on individual constructs

To investigate whether the observers understood the adjectives and used them systematically, Kendall's W was calculated separately for each construct and each of the four observer groups.

7.2.1.4.1. Familiar observers

For Stable 1 the results show that the group of first 13 constructs was understood better and used more consistently by the participants (The full list of constructs is included in Appendix 11). The first four constructs all had significant Kendall's W, while two of the last four constructs did not (Table 7.7)

Table 7.7

Consistency of familiar observers rating horses on separate constructs – Stable 1.

CONSTRUCT	W	χ^2	P
CONFIDENT-UNSURE	**0.697****	**25.105**	**0.003**
HAPPY-UNHAPPY	**0.637****	**22.943**	**0.006**
LAID-BACK-WORRIER	**0.815*****	**29.342**	**0.001**
RELIABLE-UNRELIABLE	**0.773*****	**27.817**	**0.001**
RELAXED-UPTIGHT	*0.644***	*23.190*	*0.006*
ALERT-APATHETIC	*0.453 NS*	*16.309*	*0.061*
SENSITIVE-UNFEELING	*0.267 NS*	*9.626*	*0.382*
GREEDY-NOT GREEDY	*0.605***	*21.778*	*0.001*

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Similarly familiar observers in Stable 2 also used constructs elicited by the majority most consistently (The full list of constructs with Kendall's W is included in Appendix 12). Kendall's W for the last four constructs indicates a lower degree of agreement between the observers and therefore less consistent and systematic use (Table 7.8).

Table 7.8

Consistency of familiar observers rating horses on separate constructs – Stable 2

CONSTRUCT	W	χ^2	P
BOLD – SHY	**0.592****	**23.671**	**0.009**
CLEVER – STUPID	**0.682****	**27.261**	**0.002**
SOCIABLE - ANTI-SOCIAL	**0.624****	**24.955**	**0.005**
CHEEKY – TIMID	**0.713*****	**28.537**	**0.001**
QUIET – IMPATIENT	*0.597***	*23.886*	*0.008*
GENTLE – AGGRESSIVE	*0.415 NS*	*16.619*	*0.083*
RELIABLE – UNTRUSTWORTHY	*0.393 NS*	*15.701*	*0.109*
GOOD NATURED – STROPPY	*0.503**	*20.116*	*0.028*

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

7.2.1.4.2. Experienced observers

The experienced observers demonstrated a similar pattern to familiar observers, though the differences do not appear to be large. Looking at the Kendall's W coefficients, it appears that the first four constructs had a slightly higher Kendall's W than the last two (Table 7.9). For example, the construct "intelligent-stupid" was used least consistently, while the construct "confident-nervous" appears to have been used most consistently in this group (The full list of constructs for both stables is included in Appendix 13).

Table 7.9

Consistency of experienced observers when rating horses on separate constructs Stable 1 and 2

CONSTRUCT	Stable 1 Kendall's W	Stable 2 Kendall's W
INTERESTED – DISINTERESTED	**0.405*** **	**0.156** **
BOLD – TIMID	**0.358*** **	**0.370*** **
FRIENDLY – UNFRIENDLY	**0.367*** **	**0.304*** **
CONFIDENT – NERVOUS	**0.459*** **	**0.323*** **
INTELLIGENT – STUPID	*0.109**	*0.041NS*
EASY GOING – STUBBORN	*0.261****	*0.315****
EXPERIENCED – INEXPERIENCED	*0.358****	*0.206****
PATIENT – IMPATIENT	*0.225****	*0.537****

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

7.2.1.4.3. Inexperienced observers

The inexperienced observers were not as consistent as the familiar observers. There seems to be no pattern that would differentiate Kendall's W between constructs provided by majority of the inexperienced observers and those provided by the minority (Table 7.10). The full list of results for construct consistency for inexperienced observers is included in Appendix 14.

Table 7.10

Consistency of inexperienced observers rating horses on separate constructs

CONSTRUCT	Kendall's W Stable 1	Kendall's W Stable 2
INTERESTED – DISINTERESTED	**0.427*** **	**0.301*** **
FRIENDLY – UNFRIENDLY	**0.480*** **	**0.395*** **
CALM – AGITATED	**0.197*** **	**0.654*** **
CONFIDENT – NERVOUS	**0.280*** **	**0.340*** **
SOCIABLE – UNSOCIABLE	*0.534****	*0.371****
PATIENT – IMPATIENT	*0.267****	*0.539****
GENTLE – HARSH	*0.254****	*0.547****
WELL BEHAVED – STUBBORN	*0.305****	*0.463****

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

7.2.2. Horse profiles

In this section we explore the personality profiles of horses about whom the familiar observers agreed the most and the least. We were interested to identify personality constructs associated with higher or lower observer agreement. We chose the familiar observers as they were the only group knowing the horses well.

7.2.2.1. Stable 1

Horse personality profiles show which scores observers gave to each individual horse. They also show a variation of scores on a particular construct. The familiar observers from Stable 1 rated horses on all 32 constructs (see Table 7.3, page 92, for the list of constructs). The degree of agreement was the highest for Ginger (Figure 7.11). The observers chose more extreme scores for constructs, and the horse was agreed to be more "willing to work" (construct 7), "timid" (construct 9) and "easy to lead" (construct 18). Higher degree of agreement indicates that the observers were more unanimous on deciding how to rate Ginger.

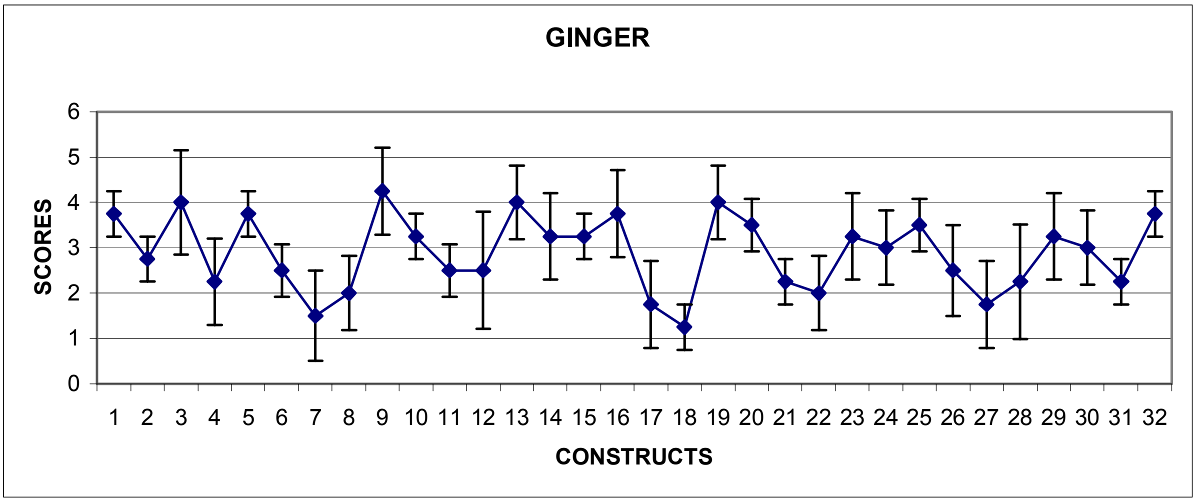


Figure 7.11: Familiar observers rating Ginger on all 32 constructs, Stable 1

Sam (Figure 7.12), on the other hand, was the only horse about whom the familiar observers did not reach a significant agreement. The scores are closer to 3, which meant that the observers could not decide easily. Standard deviation bars also indicate variation between the observers.

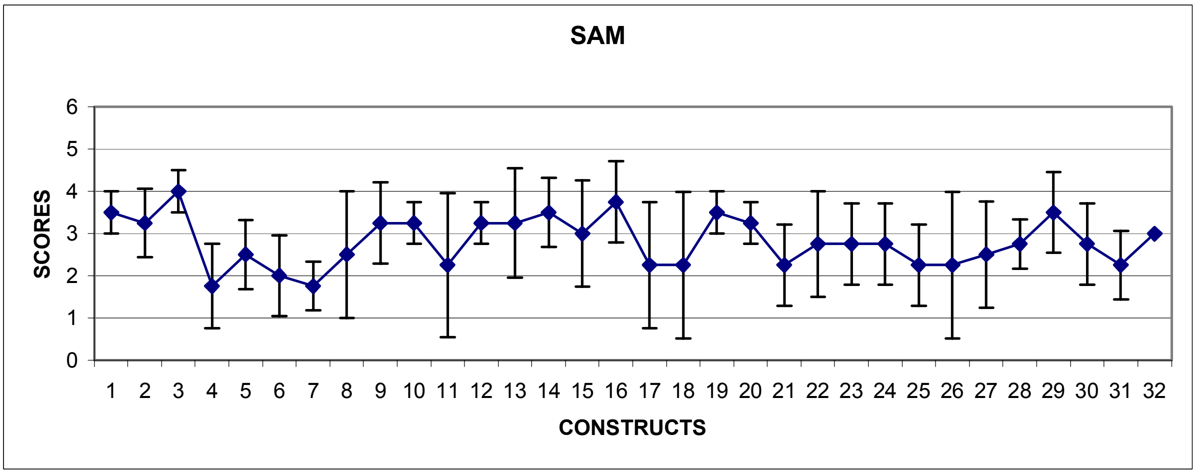


Figure 7.12: Familiar observers rating Sam on all 32 constructs, Stable 1

However, from the profile we can see that the there is not much variation between the observers for construct 3 ("laid back-worrier") and that Sam was perceived as a worrier, who was moderately unsure (construct 1), unreliable (construct 4), unwilling to work (construct 7) and a follower (constructs 16 and 19). Stable 1 was a riding school for disabled, therefore willingness to work is an important part of horses' daily routines.

7.2.2.2. Stable 2

Stable 2 on the other hand was a riding school and a livery. Again we present personality profiles of two horses: Appley had the highest agreement, and Slotty was the only horse about whom the familiar observers did not reach significant degree of agreement. Appley received more extreme scores with less variation between the observers (Figure 7.13). This means that the observers were more unanimous about the horse and they decided about the scores more easily. Appley was described as bold (construct 1), but anti-social (construct 3). He was also perceived as obedient (construct 13), easy to work with (construct 16), well mannered (construct 24) and gentle (construct 26).

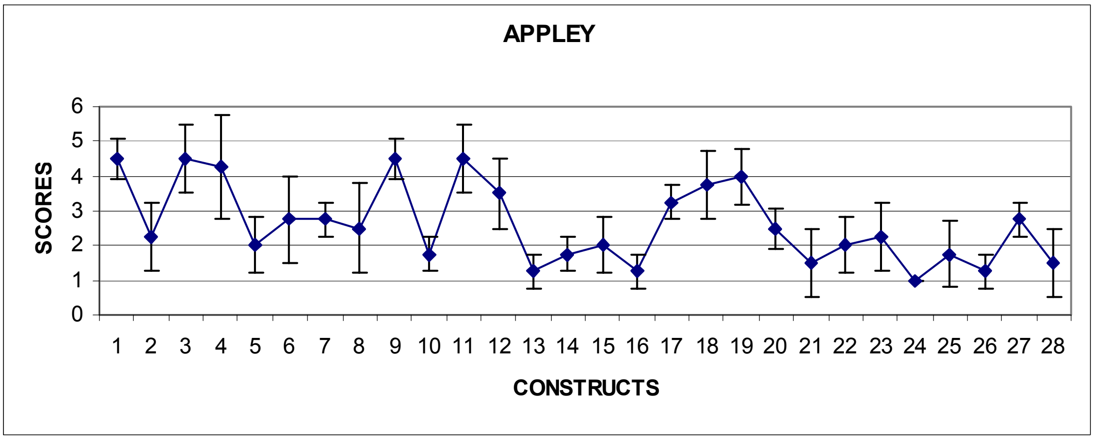


Figure 7.13: Familiar observers rating Appley on all 28 constructs, Stable 2

The degree of agreement between familiar observers was non-significant only for Slotty. His profile shows more middle scores and greater variation between the observers (Figure 7. 14).

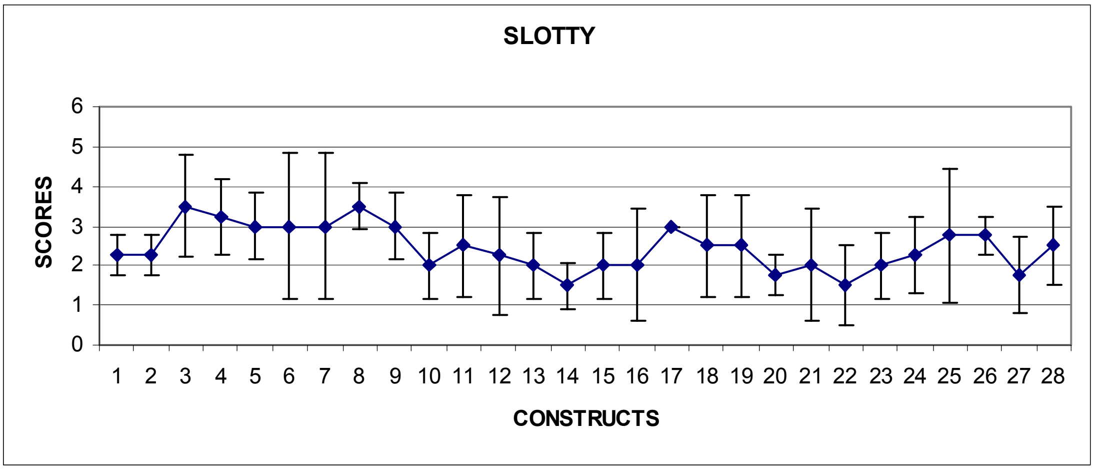


Figure 7.14: Familiar observers rating Slotty on all 28 constructs, Stable 2

Slotty was perceived as slightly aloof (construct 8), moderately interested (construct 14), but quite sensitive (construct 22) and reliable (construct 27).

7.2.3. Multidimensional scaling

In order to uncover horse personality dimensions we applied Multidimensional scaling (MDS) for both sets of data, Stable 1 and Stable 2. MDS is used to reveal the structure of a data set by calculating proximities, "which indicates how similar or how different two objects are, or are perceived to be" (Kruskal & Wish, 1978, p.7). The graphical output of MDS is a spatial configuration of points. These points in our study correspond to horses. For valid results it is recommended to have at least four horses for one dimension, nine for two dimensions, 13 for three and so on. However, even though we have conducted both two and three-dimensional models, only the two dimensional have been included in this chapter. Three dimensional models with linear regressions for identification of the dimensions are included in Appendix 15.

To determine the proportion of variance of the scaled data in the matrix that is accounted for by the corresponding distances we calculated squared correlation in distances (RSQ) for each model. The absolute value of 1 indicates that the variance has been fully explained.

7.2.3.1. Stable 1

In this section we present Euclidean distance models, based on data sets generated by all observer groups for Stable 1. Each observer group produced one perceptual map (Figures 7.15-7.17). In the following sections we will present linear regression for explanation and identification of personality dimensions. However, for presentation purposes we have labelled the dimensions on the distance models in this section.

7.2.3.1.1. Familiar observers

From the perceptual map generated by the familiar observers, Eddie was perceived as very different from Sandy or Sam. Horses Bobby, Hamish and Sam are perceived as neither high or low on the Extroversion (E) dimension. On the other hand Eddie was assessed as being extremely low on the Neuroticism (N) dimension, and Sam as extremely high. Referring to horse profiles, Sam was assessed as a moderately unsure worrier (see Figure 7.12).

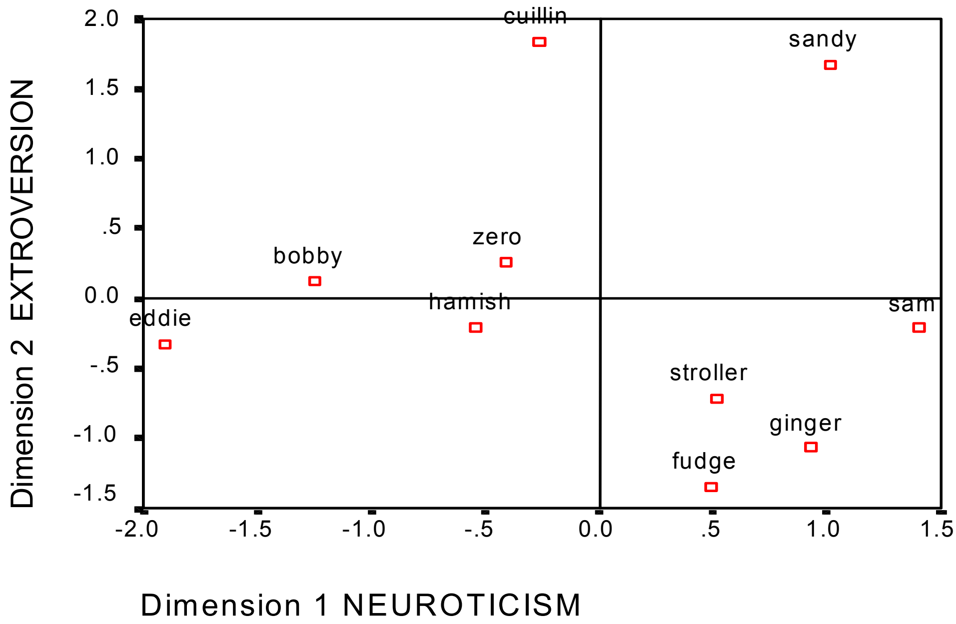


Figure 7.15

Euclidean two-dimensional distance model produced by familiar observers - Stable 1

For familiar observers from Stable 1 RSQ was 0.94608. Goodness of fit for the model was determined by calculating stress, using Kruskal's stress formula 1. Stress is the error measurement and values close to 0 indicate a good fit for the model. The stress value for familiar observers from Stable 1 was 0.09134.

7.2.3.1.2. Experienced observers

Experienced observers also perceived Eddie as quite different from Sam. However Eddie is slightly lower on E, and Sam is slightly higher on the same dimension. Hamish, however, was perceived as very interactive, interested and quite calm.

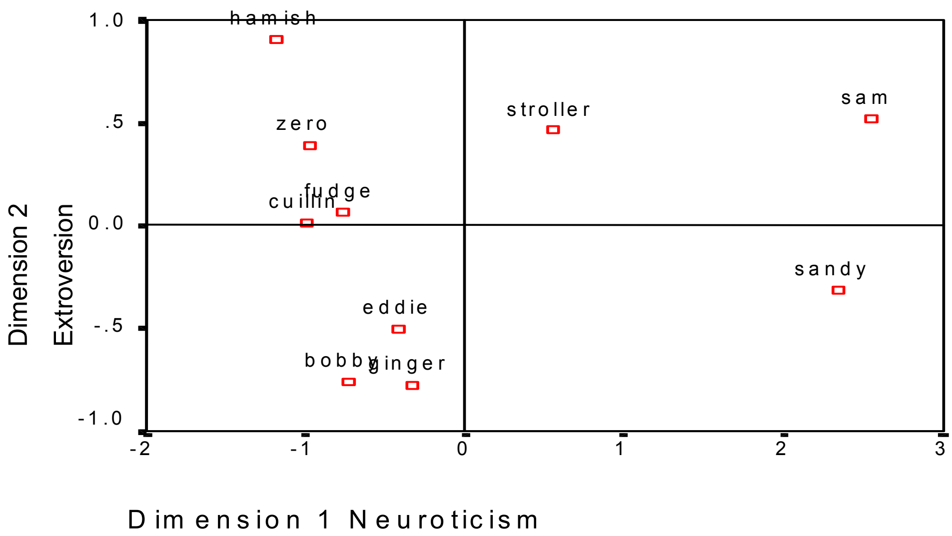


Figure 7.16

Euclidean two-dimensional distance model produced by experienced observers

The matrix for experienced observers shows good fit (stress = 0.04139), and that there was a high proportion of variance explained (RSQ = 0.99401).

7.2.3.1.3. Inexperienced observers

The neuroticism dimension in Figure 7.17 is reversed, therefore horses on the negative pole of N scored higher on this dimension. Inexperienced observers also perceived Sam as quite anxious and nervous, while Eddie was perceived as calm and slightly introverted. Hamish was positioned high on E also by inexperienced observers (Figure 7.17). It seems that experienced and inexperienced observers positioned horses in a similar order along the two dimensions.

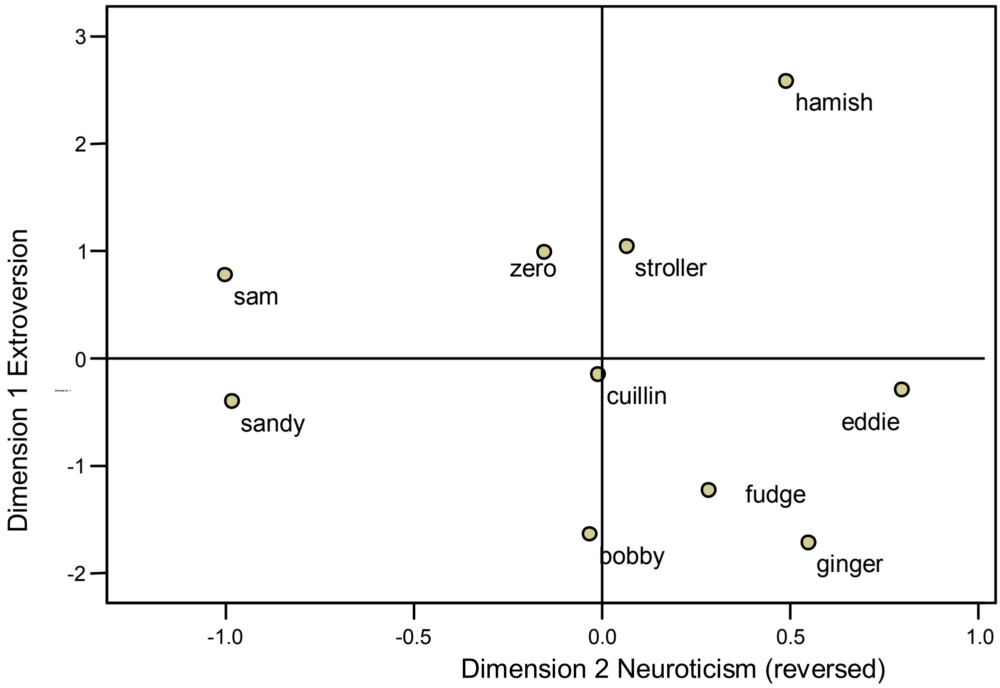


Figure 7.17
Euclidean two-dimensional distance model produced by inexperienced observers

MDS produced by inexperienced observers also indicates good fit (stress = 0.08359) with a high proportion of variance explained (RSQ = 0.96547).

7.2.3.2. Stable 2

Observers rating horses from Stable 2 also produced two and three dimensional models calculated with MDS. Below are three two dimensional models, one for each observer group (Figure 7.18-7.20. Three dimensional models for Stable 2 are included in Appendix 15.

7.2.3.2.1. Familiar observers

Reiver and Appley were perceived as the most different horses. Reiver was low on N while Appley scored the highest on the same dimension. Dolce, Rob, Stroppy and Slotty were positioned very close and therefore were not perceived in the same way as Reiver.

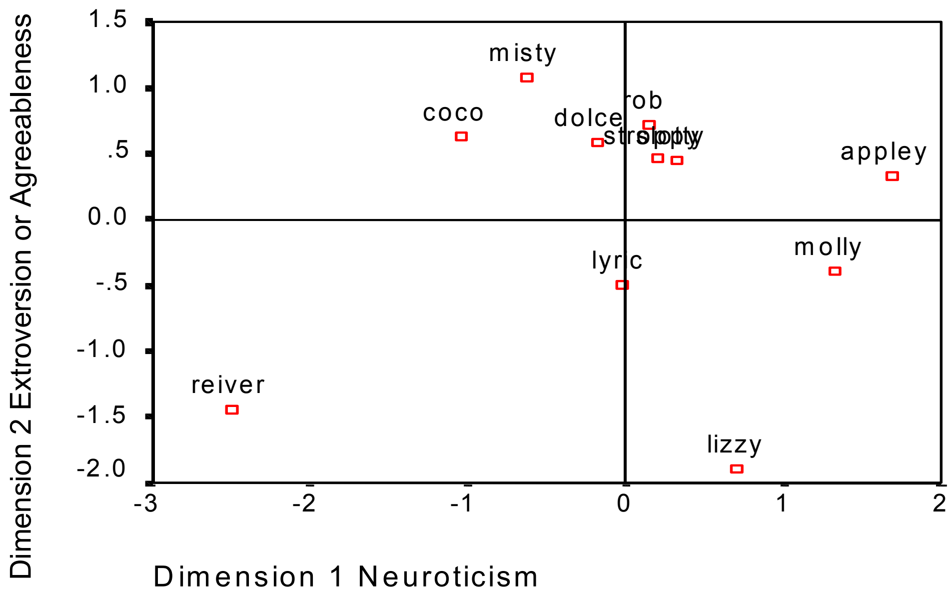


Figure 7.18

Euclidean two-dimensional distance model produced by familiar observers

For familiar observers at Stable 2 the model shows good fit (Stress = 0.06956) with a high proportion of variance explained (RSQ = 0.96884).

7.2.3.2.2. Experienced observers

Experienced observers perceived Reiver similarly to the familiar group. Appley however is positioned high on E and moderately low on N, which is quite different from familiar observers' positioning. Misty is low on E here, while familiar observers positioned her highest on the same dimension.

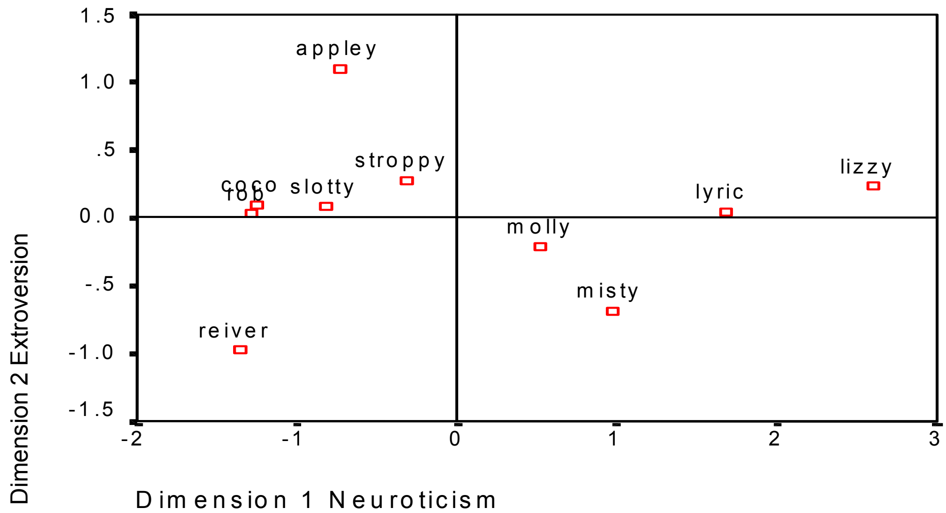


Figure 7.19

Euclidean two-dimensional distance model produced by experienced observers

Experienced observers positioned horses in a model that shows good fit (Stress = 0.05798) and a high proportion of variance is explained (RSQ = 0.98532).

7.2.3.2.3. Inexperienced observers

Inexperienced observers also positioned Appley high on E, but slightly lower on the N dimension. Reiver and Misty are both low on E, which is a similar spatial configuration to that of experienced observers.

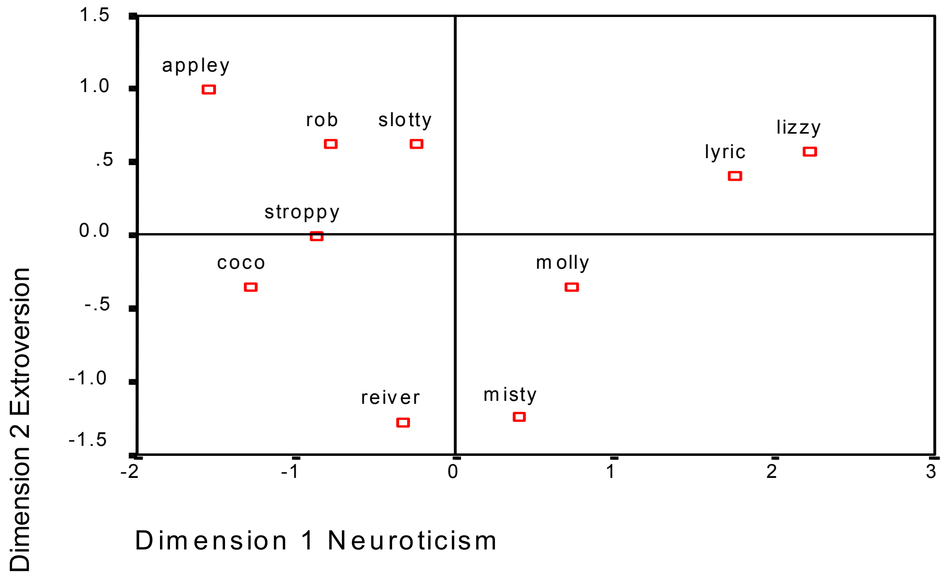


Figure 7.20
Euclidean two-dimensional distance model produced by inexperienced observers

Good fit (stress = 0.05222) and a high proportion of variance explained RSQ = 0.98421) was also found in the model calculated by the group of inexperienced observers.

7.2.3.3. MDS models for horses from both stables generated by two groups of unfamiliar observers

Both experienced and inexperienced observers rated horses by watching the video of individual horses in a human interaction test. In this section experienced (Figure

7.21) and inexperienced (Figure 7.22) observers produced perceptual maps for all the horses.

Dimension 1, which we labelled as extroversion (E) was the most obvious, however the second dimension was not that clearly defined, as we will see from the linear regression.

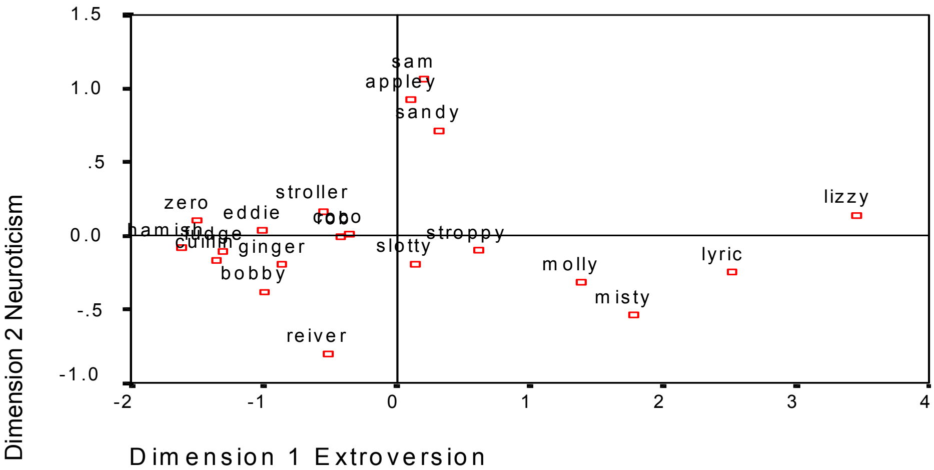


Figure 7.21

Euclidean two-dimensional distance model for both riding schools produced by experienced observers

Experienced observers positioned horses from Stable 1 lower on E scale, while inexperienced observers positioned them high on the N dimension. However, Lizzy, Lyric and Hamish seem to be perceived in a similar way by both experienced and inexperienced observers.

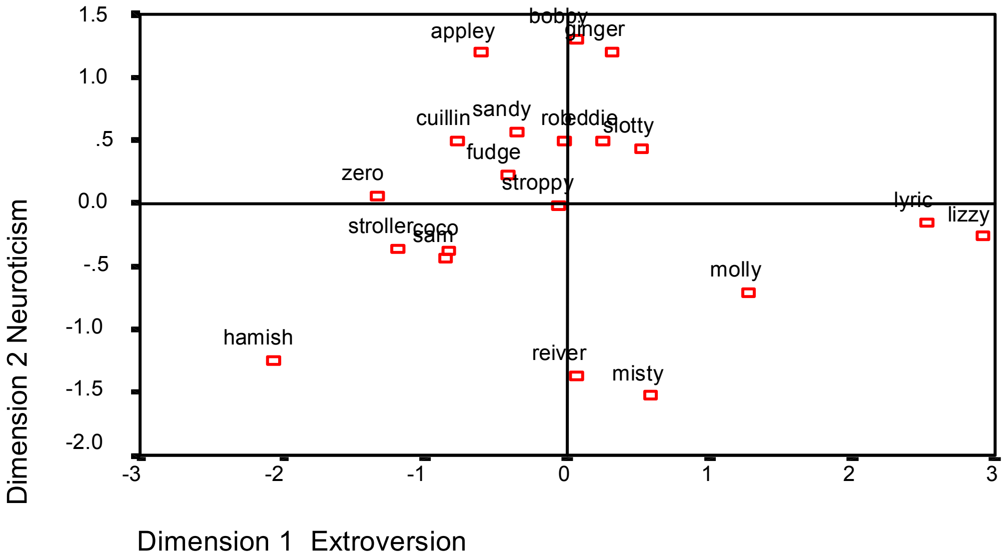


Figure 7.22

Euclidean two-dimensional distance model for both riding schools produced by inexperienced observers

The Euclidian distance model for all horses rated by the experienced observers showed good overall fit (Stress = 0.06903) and a good explanation of variance (RSQ = 0.98487).

Calculation of the Euclidian distance model for all horses rated by the inexperienced observers showed not as good fit as previous models (Stress = 0.11162). However, the squared correlation coefficient still indicated a high proportion of variance explained (RSQ = 0.94819).

7.2.4. Identification of dimensions

Identification of the dimensions discovered with MDS was the next step. To identify them we associated the mean scores for each construct or pair of adjectives with the proximity values of each dimension, using linear regression. Apart from the two dimensional models presented in this chapter we also calculated models with three dimensions. The third dimension was not clearly separated from the first two. The most meaningful third dimension was calculated for familiar observers who knew the horses well (these results are included in Appendix 15). The identified dimensions were coherent with previous horse studies (Morris et al., 2002).

7.2.4.1. Stable 1

7.2.4.1.1. Familiar Observers

First we conducted linear regression on two dimensions using all the constructs produced by familiar observers (Table 7.11).

Table 7.11

Linear regression to identify two dimensions - Stable 1 – Familiar observers

CONSTRUCT	T – Regression coefficient		Significance	
	DIM 1	DIM 2	DIM 1	DIM 2
CONFIDENT-UNSURE	**8.176*****	-2.290 NS	0.000	0.056
HAPPY-UNHAPPY	**8.806*****	-0.915 NS	0.000	0.391
LAID-BACK-WORRIER	**6.090*****	0.081 NS	0.000	0.938
RELIABLE-UNRELIABLE	2.074 NS	1.637 NS	0.077	0.146
BULLY-BULLIED	0.461 NS	-0.376	0.659	0.718
INTELLIGENT-STUPID	2.863*	**-4.375****	0.024	0.003
WILLING TO WORK-UNWILLING TO WORK	0.952 NS	-1.175	0.373	0.278
STEADY-FAST	4.656**	**6.389*****	0.002	0.000
BOLD-TIMID	4.521**	-2.443*	0.003	0.045
FRIENDLY-UNFRIENDLY	4.319**	-0.994 NS	0.003	0.353
FORWARD GOING-SLOW	-1.150 NS	**-5.672*****	0.228	0.001
DOES NOT LIKE SMALL AREAS – NOT WORRIED ABOUT IT	-1.480 NS	0.883 NS	0.182	0.406
PUSHY-DOCILE	0.198 NS	-1.719 NS	0.849	0.129

TRUSTFUL-SUSPICIOUS	**8.583*****	1.354 NS	0.000	0.218
AFFECTIONATE-ALOOF	1.762 NS	-2.179 NS	0.122	0.066
BOSSY-LIKES TO FOLLOW	2.324 NS	**-5.004****	0.053	0.002
OBEDIENT-DISOBEDIENT	1.498 NS	2.316 NS	0.178	0.054
EASY TO LEAD-STUBBORN	2.281 NS	3.286*	0.057	0.013
GOOD LEAD HORSE – PREFERS FOLLOWING	3.841**	-3.805**	0.006	0.007
INQUISITIVE-NOT INQUISITIVE	0.818 NS	**-5.957*****	0.440	0.001
RESPONSIVE-UNRESPONSIVE	0.346 NS	-2.599*	0.739	0.035
ACCEPTING-NOT ACCEPTING	**6.185*****	4.052**	0.000	0.005
SOCIABLE-LONER	0.971 NS	-1.636 NS	0.364	0.146
LIVELY-SLUGGISH	-1.781 NS	-6.141**	0.118	0.000
PLAYFUL-DULL	-2.077 NS	**-5.325*****	0.076	0.001
QUIET-EXCITABLE	3.272*	3.587**	0.014	0.009
PATIENT-IMPATIENT	3.991**	2.575*	0.005	0.037
GOOD NATURED-STROPPY	**5.951*****	0.303 NS	0.001	0.771
RELAXED-UPTIGHT	**5.533*****	0.833 NS	0.001	0.432
ALERT-APATHETIC	0.803 NS	-3.038*	0.449	0.019
SENSITIVE-UNFEELING	0.268 NS	-1.903 NS	0.796	0.099
GREEDY-NOT GREEDY	2.152 NS	-1.322 NS	0.068	0.228

P<0.05; ** P<0.01; *** P<0.001; NS - not significant.

Data from Stable 1 indicates that the first dimension is significantly correlated with the constructs "confident-unsure", "happy-unhappy", "laid back-worrier", "trustful-suspicious" and are associated with *neuroticism facets* in accordance with the Five Factor Model of personality (Costa and McCrae, 1992). The second dimension is highly associated with the constructs "steady-fast", "forward going-slow", "inquisitive-not inquisitive" and "playful-dull". The constructs could fall within either the *extroversion* or *openness* facets but also *agreeableness* (Table 7.11).

7.2.4.1.2. Experienced Observers

The data from Stable 1 produced by experienced observers indicates that the first dimension is highly correlated with the constructs "confident-nervous", "relaxed-tense", and "secure-frightened" and are correlated with *neuroticism facets* in accordance with the five factor concept of personality (Costa and McCrae, 1992). The second dimension associates significantly with the constructs "friendly-unfriendly", "sociable-unsociable", affectionate-unaffectionate", and "playful-serious". These constructs could fall within the *extroversion facets* (Table 7.12).

Table 7.12

Linear regression to identify two dimensions - Stable 1 – Experienced observers

CONSTRUCT	T – Regression coefficient		Significance	
	DIM 1 (N)	DIM 2 (E)	DIM 1	DIM 2
INTERESTED-DISINTERESTED	-3.200*	**-4.581****	0.015	0.003
BOLD-TIMID	**6.853*****	-1.126	0.000	0.297
FRIENDLY-UNFRIENDLY	-0.361	**-3.813****	0.728	0.007
CONFIDENT-NERVOUS	**14.144*****	-0.258	0.000	0.804
BORED-ALERT	5.526***	3.337*	0.001	0.012
RELAXED-TENSE	**28.099*****	-1.149	0.000	0.288
LAID BACK-SPIRITED	**11.817*****	-1.584	0.000	0.157
CALM-EXCITED	**7.950*****	-1.156	0.000	0.286
UNCONCERNED-WORRIED	**8.633*****	0.133	0.000	0.898
CONTENT-UNHAPPY	3.574**	-2.127	0.009	0.071
SOCIABLE-UNSOCIABLE	-0.668	**-8.144*****	0.525	0.000
ACTIVE-PASSIVE	-2.820*	-0.636	0.026	0.545
SECURE-FRIGHTENED	**18.753*****	**-4.018****	0.000	0.005
SPOOKY-TRUSTWORTHY	**-18.328*****	2.235	0.000	0.061
SURE-UNSURE	**18.794*****	-1.757	0.000	0.122
QUIET-ENERGETIC	**6.671*****	-0.661	0.000	0.530
TRUSTING-UNTRUSTING	4.859**	-1.957	0.002	0.091
PLACID-HOT TEMPERED	**7.207*****	-1.978	0.000	0.088
SAFE-JUMPY	**10.329*****	0.222	0.000	0.831
DOCILE-HIGHLY STRUNG	**7.559*****	0.091	0.000	0.930
SETTLED-UNSETTLED	**10.050*****	-1.710	0.000	0.131
PLAYFUL-SERIOUS	-0.415	**-3.796****	0.690	0.007
AWARE-OBLIVIOUS	-4.251**	-2.490*	0.004	0.042
GENTLE-AGGRESSIVE	0.293	-2.263	0.778	0.058
AFFECTIONATE-	0.232	**-4.104****	0.823	0.005

UNAFFECTIONATE				
STEADY-FLIGHTY	**8.514*****	-1.593	0.000	0.155
GOOD NATURED-DIFFICULT	2.071	**-3.636****	0.077	0.008
INTELLIGENT-STUPID	1.932	-1.719	0.095	0.129
EASY GOING-STUBBORN	3.537**	-0.765	0.010	0.469
EXPERIENCED-INEXPERIENCED	5.896***	-1.455	0.001	0.189
PATIENT-IMPATIENT	3.813**	-3.428*	0.007	0.011

P<0.05; ** P<0.01; *** P<0.001; NS - not significant. Degree of freedom: 9.

7.2.4.1.3. Inexperienced Observers

Data from Stable 1 produced by inexperienced observers indicates that the first dimension is highly correlated with the constructs "friendly-unfriendly", "affectionate-unaffectionate", and "sociable-unsociable". These constructs could fall within the extroversion facets. The second dimension associates significantly with the constructs "confident-nervous" and "sure-unsure" and are correlated with neuroticism facets in accordance with five factor concept of personality (Costa and McCrae, Manual) (Table 7.13).

Table 7.13

Linear regression to identify two dimensions - Stable 1 – Inexperienced observers

CONSTRUCT	T – Regression coefficient		Significance	
	DIM 1 (E/Sociability)	DIM 2 (N)	DIM 1	DIM 2
INTERESTED-DISINTERESTED	**-6.600*****	0.554	0.000	0.597
FRIENDLY-UNFRIENDLY	**-9.542*****	-0.426	0.000	0.683
CALM-AGITATED	-2.012	-1.455	0.084	0.189
CONFIDENT-NERVOUS	-2.380*	**-2.652***	0.049	0.033
HAPPY-UNHAPPY	**-10.986*****	-1.896	0.000	0.100
PLACID-HIGHLY STRUNG	-3.220**	0.211	0.015	0.839
PLAYFUL-SERIOUS	**-7.834*****	-1.719	0.000	0.129
ACTIVE-PASSIVE	-1.933	-0.931	0.095	0.383

SURE-UNSURE	-3.138*	**-2.676***	0.016	0.032
BORED-INTERESTED	**12.892*****	-0.421	0.000	0.687
DOCILE-SPIRITED	-0.222	0.557	0.830	0.595
ALERT-TIRED	-2.227	0.262	0.061	0.801
INTERACTIVE-DISTANT	**-10.310*****	-0.431	0.000	0.679
TOLERANT-INTOLERANT	**-6.469*****	-0.342	0.000	0.742
RELAXED-TENSE	-3.707**	-1.378	0.008	0.211
SECURE-INSECURE	-3.507**	-1.487	0.010	0.180
TRUSTING-DISTRUSTFUL	**-9.892*****	-1.825	0.000	0.111
EXTROVERTED –INTROVERTED	**-17.112*****	**-7.621*****	0.000	0.000
APPROACHABLE –STAND-OFFISH	**-12.714*****	-0.373	0.000	0.720
AFFECTIONATE-UNAFFECTIONATE	**-16.008*****	-1.222	0.000	0.261
BOLD-TIMID	-3.282*	-2.181	0.013	0.066
ACCEPTING-REJECTING	**-8.631*****	1.222	0.000	0.261
SOCIABLE-UNSOCIABLE	**-15.590*****	0.358	0.000	0.731
PATIENT-IMPATIENT	-3.585**	0.052	0.009	0.960
GENTLE-HARSH	-4.605**	**2.841***	0.002	0.025
WELL BEHAVED-STUBBORN	**-6.696*****	0.756	0.000	0.474

P<0.05; ** P<0.01; *** P<0.001; NS - not significant. Degree of freedom: 9.

7.2.4.1.4. Further explanation of dimensions for Stable 1

For Stable 1, familiar observers significantly associated the following constructs with the first dimension: "confident-unsure", "happy-unhappy", "laid-back-worrier", "trustful-suspicious", "accepting-not accepting", good-natured-stroppy" and "relaxed-uptight". Experienced observers associated "relaxed-tense", "sure-unsure", "secure-frightened", "confident-nervous", "laid-back-spirited", "safe-jumpy", "settled-unsettled", "unconcerned-worried", "steady-flighty" and "calm-excited" with the first dimension. In the inexperienced group association between a dimension and the following constructs e.g. "confident-nervous", "sure-unsure", "gentle-harsh" was not as pronounced as in the other two observer groups.

The descriptors produced by all observer groups and associated with the first dimension are also associated with Neuroticism facets: *Anxiety, Angry Hostility, Self-Consciousness and Impulsiveness*. Therefore we have labelled the first dimension *Neuroticism*. An overview of animal personality research (Gosling, 2001) also associated similar descriptors or traits with Neuroticism.

In the familiar group the following constructs were significantly associated with the second dimension: "steady-fast", "inquisitive-not inquisitive", "forward going-slow", "playful-dull", "bossy-likes to follow", "good lead horse-prefers following", "lively-sluggish", "quiet-excitable" and "intelligent-stupid". Experienced observers had their second dimension more clearly defined. Constructs associated with it were "sociable-unsociable", "interested-disinterested", "affectionate-unaffectionate", "friendly-unfriendly", "playful-serious" and "good natured-difficult". Inexperienced observers also generated constructs that were strongly associated with this dimension: "extroverted-introverted", "affectionate-unaffectionate", "sociable-unsociable", "bored-interested", "approachable-stand-offish", "happy-unhappy", "interactive-distant", "trusting-distrustful" and "friendly-unfriendly".

Constructs associated with the second dimension did not provide a simple solution. Some of them fall into the Intellect dimension (Gosling, 2001): "inquisitive-not inquisitive" and "intelligent-stupid". They could also be part of the *Ideas* facet of the Openness dimension (Costa and McCrae, 1992). Constructs like "bossy-likes to follow" which occur frequently with familiar observers could be part of the *Compliance* or *Modesty* facets of the Agreeableness dimension. At the same time "playful-dull", "lively-sluggish" and "quiet-excitable" are adjectives associated with the *Activity* facet of the Extroversion dimension. However, even though the dimension is not very clearly defined, we have labelled it *Extroversion*. For the experienced and inexperienced observers the constructs associated with the second dimension provided a more straightforward choice. Most of the constructs fall into the *Warmth* ("sociable-unsociable", "affectionate-unaffectionate" and "approachable-stand-offish"), but also the *Gregariousness* facet, part of the of

Extroversion dimension. Therefore, the second dimension for both unfamiliar groups was *Extroversion*.

7.2.4.2. Stable 2

7.2.4.2.1. Familiar Observers

Data from Stable 2 indicates that the first dimension is highly associated with similar constructs to those of the first dimension from Stable 1 model: "cheeky-timid", "confident-nervous", "safe-spooky", "content-unsettled" and "secure-insecure". These could be associated with *neuroticism facets*. The second dimension associates significantly with the constructs "willing to please-stubborn", "obedient-disobedient", "easy to work with-difficult", "well behaved-naughty" and "well mannered-ignorant". They could be interpreted as *agreeableness facets* (Table 7.14).

Table 7.14

Linear regression to identify two dimensions - Stable 2 – Familiar observers

CONSTRUCT	T – Regression coefficient		Significance	
	DIM 1	DIM 2	DIM 1	DIM 2
BOLD – SHY	**3.816****	-0.759 NS	0.005	0.470
CLEVER – STUPID	**3.384****	**-3.675****	0.010	0.006
SOCIABLE – ANTI-SOCIAL	0.672 NS	-1.416 NS	0.521	0.195
CHEEKY – TIMID	**4.166****	-0.815 NS	0.003	0.439
FRIENDLY – UNFRIENDLY	-0.662 NS	-1.143 NS	0.526	0.286
CALM – EXCITABLE	2.575*	-1.289 NS	0.033	0.233
LAID-BACK – HIGHLY STRUNG	2.649*	-1.558 NS	0.029	0.158
AFFECTIONATE – ALOOF	-0.796 NS	-0.780 NS	0.449	0.458
CONFIDENT – NERVOUS	**4.412****	-0.703 NS	0.002	0.502
WILLING TO PLEASE – STUBBORN	-3.098*	**-4.031****	0.015	0.004
SAFE – SPOOKY	**5.617*****	-1.975 NS	0.000	0.084

CONTENT – UNSETTLED	**5.294*****	-3.018*	0.001	0.017
OBEDIENT – DISOBEDIENT	-2.481*	**-4.940*****	0.038	0.001
INTERESTED – DISINTERESTED	-1.411 NS	-2.175 NS	0.196	0.061
LIVELY – DULL	-2.456*	0.152 NS	0.040	0.883
EASY TO WORK WITH – DIFFICULT	-1.701 NS	**-6.769*****	0.127	0.000
PLAYFUL – BORING	1.492 NS	-1.202 NS	0.174	0.264
SECURE – INSECURE	**5.115*****	-2.398*	0.001	0.043
PREDICTABLE – UNPREDICTABLE	**4.182****	**-3.518****	0.003	0.008
HAPPY – UNHAPPY	3.269*	**-5.237*****	0.011	0.001
WELL BEHAVED – NAUGHTY	-2.956*	**-8.833*****	0.018	0.000
SENSITIVE – UNRESPONSIVE	-2.847*	-2.342*	0.022	0.047
FRESH – LAZY	-2.571*	-0.273 NS	0.033	0.792
WELL MANNERED – IGNORANT	-3.231*	**-7.328*****	0.012	0.000
QUIET – IMPATIENT	-0.365 NS	-2.338*	0.724	0.048
GENTLE – AGGRESSIVE	-2.321*	-1.778 NS	0.049	0.113
RELIABLE – UNTRUSTWORTHY	3.201*	**-4.346****	0.013	0.002
GOOD NATURED – STROPPY	-2.792*	-3.014*	0.023	0.017

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

7.2.4.2.2. Experienced Observers

Data from Stable 2 produced by experienced observers indicates that the first dimension is highly associated with "relaxed-tense", "laid-back-spirited", "calm-excited" and "steady-flighty". They could be associated with *neuroticism facets* or *extroversion facets*. The second dimension associates significantly with the constructs "confident-nervous", "playful-serious", "gentle-aggressive" and could be interpreted as *extroversion facets* (Table 7.15).

Table 7.15

Linear regression to identify two dimensions - Stable 2 – Experienced observers

CONSTRUCT	T – Regression coefficient		Significance	
	DIM 1 (N)	DIM 2 (Act/E)	DIM 1	DIM 2
INTERESTED-DISINTERESTED	-0.431	-0.718	0.679	0.496
BOLD-TIMID	0.083	**6.205******	0.937	0.000
FRIENDLY-UNFRIENDLY	2.313	0.197	0.054	0.850
CONFIDENT-NERVOUS	4.913**	**6.815******	0.002	0.000
BORED-ALERT	**5.800******	1.696	0.001	0.134
RELAXED-TENSE	**16.662******	4.202**	0.000	0.004
LAID BACK-SPIRITED	**27.292******	-0.149	0.000	0.886
CALM-EXCITED	**21.100******	-0.559	0.000	0.594
UNCONCERNED-WORRIED	**8.696******	2.313	0.000	0.054
CONTENT-UNHAPPY	5.482***	1.099	0.001	0.308
SOCIABLE-UNSOCIABLE	2.129	0.684	0.071	0.516
ACTIVE-PASSIVE	**-10.376******	**5.913******	0.000	0.001
SECURE-FRIGHTENED	**9.892******	3.402*	0.000	0.011
SPOOKY-TRUSTWORTHY	**-11.495******	-0.435	0.000	0.677
SURE-UNSURE	**6.877******	**5.579******	0.000	0.001
QUIET-ENERGETIC	**16.688******	-4.772**	0.000	0.002
TRUSTING-UNTRUSTING	**6.909******	3.114*	0.000	0.017
PLACID-HOT TEMPERED	**21.356******	-3.988**	0.000	0.005
SAFE-JUMPY	**11.587******	-0.334	0.000	0.748
DOCILE-HIGHLY STRUNG	**21.858******	-2.057	0.000	0.079
SETTLED-UNSETTLED	**10.716******	0.780	0.000	0.461
PLAYFUL-SERIOUS	-0.809	**3.174***	0.445	0.016
AWARE-OBLIVIOUS	-1.846	-1.513	0.107	0.174
GENTLE-AGGRESSIVE	**12.975******	**-7.089******	0.000	0.000
AFFECTIONATE-UNAFFECTIONATE	2.333	-0.339	0.052	0.745
STEADY-FLIGHTY	**25.952******	-1.899	0.000	0.099
GOOD NATURED-DIFFICULT	**9.204******	-1.784	0.000	0.118
INTELLIGENT-STUPID	0.435	-1.937	0.677	0.094
EASY GOING-STUBBORN	**6.264******	-2.497*	0.000	0.041
EXPERIENCED-INEXPERIENCED	**7.694******	1.418	0.000	0.199
PATIENT-IMPATIENT	**10.358******	**-3.412***	0.000	0.011

P<0.05; ** P<0.01; *** P<0.001; NS - not significant.

7.2.4.2.3. Inexperienced Observers

Data from Stable 2 indicates that the first dimension is highly associated with "calm-agitated", "placid-highly strung", "docile-spirited", and "tolerant-intolerant". They could be associated with *neuroticism or agreeableness facets*. The second dimension associates significantly with the constructs "friendly-unfriendly", "interactive-distant", and "sociable-unsociable". They could be associated with *extroversion facets* (Table 7.21).

Table 7.16

Linear regression to identify two dimensions - Stable 2 – Inexperienced observers

CONSTRUCT	T – Regression coefficient		Significance	
	DIM 1 (N)	DIM 2 (E)	DIM 1	DIM 2
INTERESTED-DISINTERESTED	0.483	**3.868****	0.644	0.006
FRIENDLY-UNFRIENDLY	4.697**	**4.097****	0.002	0.005
CALM-AGITATED	**15.499*****	-0.961	0.000	0.369
CONFIDENT-NERVOUS	4.559**	3.215*	0.003	0.015
HAPPY-UNHAPPY	**7.535*****	**5.376*****	0.000	0.001
PLACID-HIGHLY STRUNG	**33.730*****	-2.699*	0.000	0.031
PLAYFUL-SERIOUS	-0.692	**4.391****	0.511	0.003
ACTIVE-PASSIVE	-7.617***	**5.220*****	0.000	0.001
SURE-UNSURE	2.192	2.675*	0.064	0.032
BORED-INTERESTED	0.673	-2.859*	0.523	0.024
DOCILE-SPIRITED	**31.033*****	**-10.012*****	0.000	0.000
ALERT-TIRED	-4.083**	2.159	0.005	0.068
INTERACTIVE-DISTANT	0.277	**5.417*****	0.790	0.001
TOLERANT-INTOLERANT	**18.244*****	**5.587*****	0.000	0.001
RELAXED-TENSE	**14.391*****	3.127*	0.000	0.017
SECURE-INSECURE	**7.914*****	3.797**	0.000	0.007
TRUSTING-DISTRUSTFUL	**11.844*****	**6.127*****	0.000	0.000
EXTROVERTED –INTROVERTED	-4.811**	**5.446*****	0.002	0.001
APPROACHABLE –STAND-OFFISH	**7.826*****	**6.291*****	0.000	0.000
AFFECTIONATE-UNAFFECTIONATE	3.873**	3.861**	0.006	0.006

BOLD-TIMID	-3.657**	4.152**	0.008	0.004
ACCEPTING-REJECTING	**6.916*** **	4.356**	0.000	0.003
SOCIABLE-UNSOCIABLE	4.703**	**8.380*** **	0.002	0.000
PATIENT-IMPATIENT	**13.540*** **	-1.005	0.000	0.348
GENTLE-HARSH	**9.920*** **	-0.634	0.000	0.546
WELL BEHAVED-STUBBORN	**13.037*** **	-0.556	0.000	0.595

P<0.05; ** P<0.01; *** P<0.001; NS - not significant.

7.2.4.2.4. Further explanation on identification of the dimensions for Stable 2

Familiar observers from the second Stable also generated their own scoring list. Constructs associated with the first dimension were "safe-spooky", "content-unsettled", "secure-insecure", "confident-nervous", "predictable-unpredictable" and "bold-shy". They could be associated with Neuroticism facets like Anxiety and Self-consciousness, but also with the Extroversion facet, *Assertiveness*. However, we thought the dimension the constructs fit best is Neuroticism. In the unfamiliar observer groups constructs associated with the first dimension corresponded to traits grouped under Neuroticism facets (Costa and McCrae, 1992). Experienced observers associated "laid-back-spirited", "steady-flighty", "docile-highly strung", "placid-hot tempered" and "calm-excited" with the first dimension. Inexperienced observers had similar constructs linked to the first dimension: "placid-highly strung", "docile-spirited", "tolerant-intolerant", "calm-agitated" and "relaxed-tense". Inexperienced observers also associated "patient-impatient" and "well behaved-stubborn" with the first dimension.

Overall, constructs grouped around the first dimension for both unfamiliar groups seemed to define Neuroticism. However, the last two constructs provided by the inexperienced observers could be grouped under the *Compliance* facet of Agreeableness.

The second dimension provided by the familiar observers was Agreeableness and constructs significantly correlated with it were "well behaved-naughty", "well mannered-ignorant", "easy to work with-difficult", "happy-unhappy", "obedient-disobedient" and "reliable-untrustworthy". Constructs associated with the second dimension for experienced observers were "gentle-aggressive", "confident-nervous", "bold-timid" and "active-passive". These adjectives are grouped under Extroversion facets: Assertiveness and Activity. Inexperienced observers generated the following constructs significantly correlated with the second dimension: "docile-spirited", "sociable-unsociable", "approachable-stand-offish", "trusting-distrustful" and "interactive-distant". They fit facets of *Extroversion*.

7.2.4.3. Stable 1 and Stable 2 on the same MDS map

7.2.4.3.1. Experienced Observers

Data from both stables indicate that the first dimension is highly associated with "relaxed-tense", "unconcerned-worried", "laid back-spirited", "docile-highly strung", and "steady-flighty". They could be associated with *extroversion facets*. The second dimension associates significantly with the constructs "confident-nervous", relaxed-tense", "sure-unsure", and "gentle-aggressive". They could be associated with *neuroticism or extroversion facets* (Table 7.17). However, the distinction between the two dimensions cannot be made very clearly.

Table 7.17

Linear regression to identify two dimensions - Stable 1 and Stable 2 – Familiar observers

CONSTRUCT	T – Regression coefficient		Significance	
	DIM 1 (E/Act)	DIM 2 (N/E)	DIM 1	DIM 2
INTERESTED-DISINTERESTED	-1.365	-2.150*	0.190	0.046
BOLD-TIMID	2.366*	**6.068*** **	0.030	0.000
FRIENDLY-UNFRIENDLY	3.016**	-1.725	0.008	0.103
CONFIDENT-NERVOUS	**9.240*** **	**8.197*** **	0.000	0.000
BORED-ALERT	**8.241*** **	2.037	0.000	0.058
RELAXED-TENSE	**23.938*** **	**7.513*** **	0.000	0.000
LAID BACK-SPIRITED	**33.480*** **	-1.302	0.000	0.210
CALM-EXCITED	**25.129*** **	-1.418	0.000	0.174
UNCONCERNED-WORRIED	**16.179*** **	**4.491*** **	0.000	0.000
CONTENT-UNHAPPY	**8.747*** **	0.559	0.000	0.583
SOCIABLE-UNSOCIABLE	2.285*	-1.265	0.035	0.223
ACTIVE-PASSIVE	**-10.411*** **	**3.442** **	0.000	0.003
SECURE-FRIGHTENED	**17.618*** **	**5.623*** **	0.000	0.000
SPOOKY-TRUSTWORTHY	**-18.156*** **	-2.993**	0.000	0.008
SURE-UNSURE	**12.489*** **	**6.805*** **	0.000	0.000
QUIET-ENERGETIC	**24.065*** **	**-6.119*** **	0.000	0.000
TRUSTING-UNTRUSTING	**9.043*** **	1.851	0.000	0.082
PLACID-HOT TEMPERED	**36.556*** **	**-6.264*** **	0.000	0.000
SAFE-JUMPY	**20.399*** **	0.618	0.000	0.545
DOCILE-HIGHLY STRUNG	**27.631*** **	**-3.103** **	0.000	0.006
SETTLED-UNSETTLED	**19.034*** **	2.112*	0.000	0.050
PLAYFUL-SERIOUS	-1.131	1.044	0.274	0.311
AWARE-OBLIVIOUS	-4.434***	-2.561*	0.000	0.020
GENTLE-AGGRESSIVE	**15.568*** **	**-8.080*** **	0.000	0.000
AFFECTIONATE-UNAFFECTIONATE	3.077**	-1.815	0.007	0.087
STEADY-FLIGHTY	**31.549*** **	-2.538*	0.000	0.021
GOOD NATURED-DIFFICULT	**11.178*** **	**-3.040** **	0.000	0.007
INTELLIGENT-STUPID	-0.753	0.331	0.462	0.744
EASY GOING-STUBBORN	**6.556*** **	0.019	0.000	0.985
EXPERIENCED-INEXPERIENCED	**10.987*** **	1.699	0.000	0.108
PATIENT-IMPATIENT	**16.206*** **	**-4.493*** **	0.000	0.000

P<0.05; ** P<0.01; *** P<0.001; NS - not significant. Degree of freedom: 9.

7.2.4.3.2. Inexperienced Observers

Data from both stables indicate that the first dimension is highly associated with "placid-highly strung", "docile-spirited", "trusting-distrustful", and "patient-impatient". These could be associated with *extroversion (activity) or neuroticism facets*. The second dimension associates significantly with the constructs "interested-disinterested", "friendly-unfriendly", "active-passive", and "interactive-distant". They could be associated with *extroversion facets* (Table 7.18). However, the distinction between the two dimensions again cannot be made very clearly.

Table 7.18

Linear regression to identify two dimensions – Stable 1 and Stable 2 – Inexperienced observers

CONSTRUCT	T – Regression coefficient		Significance	
	DIM 1 (E/Act)	DIM 2 (N/E)	DIM 1	DIM 2
INTERESTED-DISINTERESTED	3.977**	**7.303*** **	0.001	0.000
FRIENDLY-UNFRIENDLY	**10.638*** **	**7.398*** **	0.000	0.000
CALM-AGITATED	**14.331*** **	-5.210***	0.000	0.000
CONFIDENT-NERVOUS	5.311***	0.929	0.000	0.366
HAPPY-UNHAPPY	**14.197*** **	**6.077*** **	0.000	0.000
PLACID-HIGHLY STRUNG	**36.342*** **	**-11.574*** **	0.000	0.000
PLAYFUL-SERIOUS	1.991	**6.524*** **	0.063	0.000
ACTIVE-PASSIVE	**-10.680*** **	**11.034*** **	0.000	0.000
SURE-UNSURE	4.462***	1.469	0.000	0.160
BORED-INTERESTED	-1.452	**-7.109*** **	0.165	0.000
DOCILE-SPIRITED	**23.747*** **	**-14.998*** **	0.000	0.000
ALERT-TIRED	-4.258***	5.000***	0.001	0.000
INTERACTIVE-DISTANT	4.724***	**11.098*** **	0.000	0.000
TOLERANT-INTOLERANT	**16.123*** **	1.754	0.000	0.097
RELAXED-TENSE	**15.045*** **	-1.176	0.000	0.256
SECURE-INSECURE	**9.284*** **	1.237	0.000	0.233
TRUSTING-DISTRUSTFUL	**20.569*** **	5.584***	0.000	0.000
EXTROVERTED –INTROVERTED	-2.935**	**8.465*** **	0.009	0.000

APPROACHABLE –STAND-OFFISH	**16.025*****	**9.958*****	0.000	0.000
AFFECTIONATE-UNAFFECTIONATE	**11.241*****	**7.769*****	0.000	0.000
BOLD-TIMID	-2.488*	**6.537*****	0.024	0.000
ACCEPTING-REJECTING	**11.181*****	5.730***	0.000	0.000
SOCIABLE-UNSOCIABLE	**14.507*****	**14.352*****	0.000	0.000
PATIENT-IMPATIENT	**20.618*****	-4.759***	0.000	0.000
GENTLE-HARSH	**12.030*****	-2.513*	0.000	0.022
WELL BEHAVED-STUBBORN	**12.392*****	-0.542	0.000	0.595

P<0.05; ** P<0.01; *** P<0.001; NS - not significant. Degree of freedom: 9.

7.2.5. Correlation of MDS dimensions between three sets of observers for Stable 1 and Stable 2

7.2.5.1. Stable 1

We were interested in cross validation of personality rating across the observer groups. The familiar, experienced and inexperienced observers rated the horses and positioned them along two dimensions. The Spearman rho correlation coefficient showed highly significant correlation for Extroversion (rho=0.952; P<0.001) between experienced and inexperienced observers. However, there was no significant correlation for dimension 1, which we labelled *Neuroticism*. The Spearman rho also showed significant correlation between familiar and inexperienced observers, when they were rating horses on constructs associated with Dimension 1, Neuroticism (rho = 0.818; P<0.01). There was no significant correlation between familiar and experienced observers on N (Table 7.19).

Table 7.19

Spearman rho - Correlation of dimensions between three sets of observers – Stable 1

Observers	Experienced	Inexperienced
Familiar	Dim1 Neuroticism 0.491NS	Dim1 Neuroticism 0.818**
	Dim2 Extroversion/Openness 0.127NS	Dim2 0.018NS
Experienced		Dim1 Neuroticism 0.418NS
		Dim2 Extroversion 0.952***

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

7.2.5.2. *Stable 2*

When rating horses from Stable 2 there was significant correlation between horse personality scores provided by experienced and inexperienced observers. The correlation was significant for both dimensions, neuroticism (rho=0.685, p<0.05), and extroversion (rho= 0.661, p<0.05) (Table 7.20).

Table 7.20

Spearman rho - Correlation of dimensions between three sets of observers – Stable 2

Observers	Experienced	Inexperienced
Familiar	Dim1 Neuroticism 0.382NS	Dim1 Neuroticism 0.067NS
	Dim2 Extroversion/Openness 0.055NS	Dim2 Extroversion/Openness 0.030NS
Experienced		Dim1 Neuroticism 0.685*
		Dim2 Extroversion 0.661*

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

7.2.5.3. *Stable 1 and Stable 2, Video observers*

Finally, we were interested to see if there was a correlation between the personality scores when we put them together and treated horses from both stable as one group. Even though the experienced and inexperienced observers had their own scoring lists, they were applied to both stables. Experienced and inexperienced observers rated the horses from both stables. The Spearman correlation coefficient was highly significant (rho=0.626, p<0.01) for dimension 1, which we labelled as Extroversion, but not for dimension 2, labelled as neuroticism. We could see that the second dimension was not well defined (Table 7.21).

Table 7.21

Spearman rho - Correlation of dimensions between two sets of observers – Stable 1 and 2

Observers	Inexperienced
Experienced	Dim1 Extroversion -0.626**
	Dim2 Neuroticism 0.162NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

7.3. Synopsis

In this chapter we have presented the results from our first study which involved two small stables and three separate groups of observers. There were 21 horses and 45 observers in this study. The horses were aged six to 24 years. The participants were all women aged 16 to 56. There were three groups of observers: familiar (who knew their own horses well) and unfamiliar, further divided into two groups: experienced (who had had previous experience with horses but did not know the horses in our study) and inexperienced (who had had no experience with horses). Familiar observers rated their horses on the basis of their previous experience with them, unfamiliar ones watched the videos of the horses in a human interaction test.

The method of data collection was Repertory grid technique which allowed the observers to generate their own constructs and rate the horses on them. We were interested in the extent to which the observers agreed when they rated the horses on various constructs describing the quality of their behaviour. The Kendall's coefficient of concordance W showed a significant degree of agreement in most cases in all three observer groups. Horse personality profiles were a graphical representation of the mean scores each horse received from all the observers. Some horses were perceived as more shy, interested or friendly.

In order to find out personality dimensions in horses have, we employed Multidimensional scaling which produced two clear dimensions. Linear regression analysis identified these as *Neuroticism* and *Extroversion*.

Finally we were interested in cross validation of personality scores for horses across the observer groups. The Spearman rho correlation coefficient showed little or no correlation between familiar and unfamiliar observers.

CHAPTER 8

Horse Personality – Study 2

8.1. Methods

The design of the second study includes the improvements made from the first study. There was no indication from the first study that having knowledge about horses in general would significantly influence the degree of agreement. Therefore there were only two observer groups this time: observers familiar with the horses and those unfamiliar. Familiar observers worked with the horses and knew them well. Unfamiliar observers did not know the horses but some of them had had previous experience with horses.

In this study we increased the number of horses and familiar observers. The number of unfamiliar observers was proportional to the number of familiar observers per stable. As there was no consistent correlation between horse personality scores provided by familiar and unfamiliar observers in Study 1, we decided to provide unfamiliar observers with an additional standardised novel object test. This was an open umbrella test, which was added to the video so that the unfamiliar observers had more information about the horses. The open umbrella test has been used previously with horses (Visser et al., 2003). Our hypothesis is that introducing a more traditional standardised test will influence the correlation between the two observer groups. The structure of this study is illustrated in Figure 7.2 in the previous Chapter.

8.1.1. Animals and housing

There were 38 horses from three different stables in this study. We started the experiment with five stables, but two of them withdrew before the data collection was finished. The horses from the three remaining stables varied in age and breed. They were from 3 to 21 years, and various breeds. Thoroughbreds were largely represented. There were 23 geldings and 15 mares (tables 8.1-8.3). The horses were usually kept in separate pens, unless they were taken out for inside or outside exercise or riding. The sex and age distribution of the horses is shown in the tables 8.1-8.3.

For the first part of the experiment in which horses were rated by the familiar raters, there were no specific settings for horses. In the second part of the experiment 5-minute video clips of horses interacting with a human interactor were made. The procedure of recording the videos will be explained later in the chapter.

Table 8.1

Horse information in Stable 1

HORSE	AGE–years (2002)	SEX	BREED
RYAN	10	Gelding	Thoroughbred x
JACK	15	Gelding	Thoroughbred
BARNEY	13	Gelding	Thoroughbred x
DICE	10	Gelding	Cob
BLUE	8	Gelding	Thoroughbred x
TIA	4	Mare	Thoroughbred x

RHUM	5	Gelding	WBx
DANNY	10	Gelding	Thoroughbred
DIZZY	9	Gelding	WB
ZAK	8	Gelding	Dales x
TICA	21	Mare	Thoroughbred x
GUNTHER	20	Gelding	Thoroughbred x
DUKE	11	Gelding	Thoroughbred x

Table 8.2

Horse information in Stable 2

HORSE	**AGE–years (2002)**	**SEX**	**BREED**
MILLIE	5	Mare	Welsh Mountain
HARRY	4	Gelding	NewBrest
JAZZ	8	Gelding	Welsh x
FLASH	3	Gelding	NewBrest
LIL	5	Mare	Welsh Mountain
QUENTIN	4	Gelding	Welsh cob
MR BEAN	8	Mare	Welsh x
MISTLETOE	3	Gelding	Thoroughbred x
KURT	3	Gelding	Welsh
MOLLY	7	Mare	Welsh x Connemara
ROCKY	3	Gelding	Native pony
WILLOW	8	Mare	Welsh
BLOSSOM	4	Mare	Coloured cob

Table 8.3

Horse information in Stable 3

HORSE	**AGE–years (2001)**	**SEX**	**BREED**
BRUNO	9	Gelding	Sec. D.
SOPHIE	7	Mare	Thoroughbred
FLEUR	4	Mare	Thoroughbred x
GANDALF	7	Gelding	Arab
JESTER	11	Mare	Thoroughbred
DILLY	6	Mare	Thoroughbred

CORK	14	Gelding	Native x
PETRA	5	Mare	Arab
DANNY	9	Gelding	Cob
CUILLEN	4	Gelding	Exmoor
PANDORA	5	Mare	Welsh x
BESS	6	Mare	Thoroughbred

8.1.2. The observers

Two groups of observers, familiar and unfamiliar, participated in this study. All observers were native English speakers. The familiar observers were familiar with the horses they rated. The unfamiliar participants did not know the horses at all. All familiar observers were women. To eliminate variation due to sex differences we also recruited female unfamiliar participants.

8.1.2.1. Familiar observers

There were 30 familiar observers who started the experiment, 24 completed the experiment: seven observers in Stable 1, eight observers in Stable 2 and nine observers in Stable 3. The raters were all women aged 17 to 44, mean age 26.5 years. They were working either as trainers or grooms. All the observers had observed and interacted with the horses in a variety of situations and they had known the animals for periods of five months to 15 years.

8.1.2.2. Unfamiliar observers

Seventeen unfamiliar observers, all psychology students at an Edinburgh college originally started the experiment. Ten completed all three sessions and filled in all the questionnaires. They all observed horses from both stables. Seven observers withdrew from the study. The age of the participants varied from 17 to 32, with a mean age of 20.8. They were all female.

8.1.3. Experimental procedure

The study consisted of two separate parts. In the first part the horses were rated by familiar observers; in the second part the unfamiliar participants rated the horses by watching their individual interaction with a human interactor recorded on video. The videos were one minute and 30 seconds long for the elicitation procedure and two minutes 40 seconds for the scoring. Scoring videos included a one minute blank screen between films of individual horses, which gave the observers time for rating.

8.1.3.1. Collecting information about the observers

As in the first study all the observers completed a NEO personality test (Costa & McCrae, 1992) and a 33 item Emotional intelligence (EI) test (Schutte et al, 1998).

8.1.3.2. Repertory Grid Technique

In this study we used a modified Repertory grid technique (RGT), which was described in detail in Chapters 6 and 7. Elicitations and interviews for participants to elaborate and explain the constructs were done individually. The observers were also given printed instructions for scoring (These are included in Appendix 18).

8.1.3.2.1. Rating by the familiar observers

The procedure was identical to the one we used in the first study. The introduction and instructions were delivered to one observer group at a time. The elicitation procedure was conducted individually and it took up to one and a half hours per person. The interviews were also done individually and the participants were asked to identify identical constructs.

The constructs for each stable separately were typed up with a clear indication of how many participants used the constructs. Initially the first pole was determined, afterwards the opposites were identified and the most frequent ones were retained as the pair in the bipolar construct. If the opposite had been used already the next most frequent one was retained. Some descriptors were generated by participants from all five stables. All 30 familiar observers from five unrelated stables were treated as one group and the most frequently occurring descriptors were placed on the uniform scoring list. To produce the uniform scoring list we simply counted the number of participants who came up with a descriptor and the number of participants who came up with the same opposites. The order of the constructs on the scoring list was chosen according to two criteria: first, the most frequently occurring descriptors were put in descending order with the most frequently occurring opposites. For example: "bossy-submissive" was placed before "happy-unhappy" because there were only nine participants who decided to put unhappy as an opposite to happy and 10 participants who decided that submissive was an opposite to bossy (Table 8.4). Similarly to the procedures in the first study, highly idiosyncratic constructs or repetitions were eliminated. The full list of constructs provided by the familiar observers with the number of participants who generated the constructs are included in Appendix 16.

Table 8.4

Example of bipolar constructs

BOSSY	16	SUBMISSIVE	**10**
HAPPY	16	UNHAPPY	**9**

The final scoring list included 34 bipolar constructs out of 72 produced by the familiar observers. At least a quarter of participants (seven out of 30) had to come up with the first part of the descriptor to be included on the scoring list. Those 34 constructs represent a list of behavioural descriptors, the most frequently used by people who work with horses (Table 8.5). The complete uniform scoring list for familiar observers is included in Appendix 17.

Table 8.5

List of constructs generated by the familiar observers

1. FRIENDLY	UNFRIENDLY
2. ATTENTION SEEKER	INDEPENDENT
3. CONFIDENT	TIMID
4. BOLD	SHY
5. EASY TO HANDLE	DIFFICULT
6. SOCIABLE	UNSOCIABLE
7. BOSSY	SUBMISSIVE
8. HAPPY	UNHAPPY
9. LAID BACK	HIGHLY STRUNG
10. CALM	EXCITABLE
11. PLAYFUL	BORING
12. AFFECTIONATE	AGGRESSIVE
13. OBEDIENT	DISOBEDIENT
14. NERVOUS	BOLD
15. INTERESTED	DISINTERESTED
16. AGGRESSIVE	GENTLE
17. DOMINANT	SUBMISSIVE
18. GENTLE	ROUGH
19. EXPERIENCED	INEXPERIENCED
20. MATURE	IMMATURE
21. EASY TO WORK WITH	DIFFICULT
22. CHEEKY	WELL BEHAVED
23. SECURE	INSECURE
24. NASTY	NICE
25. BRAVE	SCARED
26. FORWARD	HESITANT
27. BULLY	BULLIED
28. STUBBORN	WILLING
29. INTELLIGENT	THICK
30. RELAXED	TENSE

31. CONTENT	GRUMPY
32. PATIENT	IMPATIENT
33. CLEVER	STUPID
34. BARGY	QUIET

Due to different work dynamic in the stables some participants scored the horses in the presence of the experimenter and some did it in their own time. For that reason the instructions were included as the first page of the scoring pack (The instructions are included in Appendix 18). The participants had to rate their own horses on the 5-point scale per construct, having one pole of the constructs on one side and the second pole on the other side. They were told to circle the appropriate number on the scale. If they thought that a horse was very friendly they had to circle 1, if they thought it was moderately friendly they circled 2, if they thought it was neither friendly nor unfriendly they circled 3 and so on. They were asked to think about the horses' behaviour across different situations, and score them in accordance to their interactions with other horses, people and on the basis of their own experience with them. All constructs were kept and used for the analysis.

8.1.3.2.2. Rating by the unfamiliar raters - Video part

Initially we decided to record the video of horses in both a human interaction test and in a novel object test at three stables. One of the stables decided to withdraw from the study so we recorded and edited videos of horses from Stables 1 and 2.

In contrast to the first study we added a novel object test – an open umbrella test. The procedure was the same in all three stables. Testing took place in one day, with a duration of 6 hours per stable. The horses were not previously trained for the experiment. All the videos were made in an outside paddock. The outside paddock was not the place where the horses were usually kept, which meant that the locations were fairly new to them. At both stables we restricted the paddock with a rope and posts to make an area of approximately 10 x 5 m. A human interactor who was not familiar with the horses was standing within the restricted area. Horses were brought into the area by a familiar person and released as soon as they entered the area. The procedure was the same as in Study 1 (see section 7.1.3.1.2.). After 7 minutes the interactor left the paddock and placed an open umbrella within the restricted area. Horses' response to the umbrella was recorded for an additional 3 minutes per horse.

The elicitation procedure was carried out as described in section 7.1.3.1.2.1. However, the one-minute inserts of each horse interacting with a human interactor was followed by 30 seconds of the open umbrella test. We used a dyadic elicitation procedure comparing pairs of horses. The difference was that the horses from Stables 1 and 2 were shown interchangeably in pairs for better contrast and easier generation of the constructs. The elicitation this time was done in a group and it took approximately 1 hour to complete the session. The individual interviews were replaced by a request to elaborate each construct and identify the identical constructs made by the experimenter at the end of the elicitation part. Having done the elicitation procedure the experimenter composed the uniform list of constructs based on individually elicited pairs of adjectives. The scoring form produced by the unfamiliar observers contained 28 items (Table 8.6). The constructs were put in order in accordance with the number of participants who used the constructs. For example, constructs common to everybody (interested-uninterested) were identified and put at the beginning of the scoring list. Highly idiosyncratic constructs or repetitions were eliminated.

Table 8.6

Scoring list produced by unfamiliar observers

1. INTERESTED	UNINTERESTED
2. BOLD	SHY
3. CURIOUS	INDIFFERENT
4. PLAYFUL	BORED
5. OUTGOING	TIMID
6. CALM	NERVOUS
7. HAPPY	SAD
8. ENTERTAINED	BORED
9. AT EASE	SCARED
10. INQUISITIVE	DISINTERESTED
11. CONFIDENT	WARY
12. FRIENDLY	UNFRIENDLY
13. BRAVE	SCARED
14. LAID-BACK	WORRIED
15. CALM	BOISTEROUS
16. AFFECTIONATE	COLD
17. STEADY	JUMPY
18. NOSY	NOT BOTHERED
19. STRONG	ANXIOUS
20. DOMINANT	SUBMISSIVE
21. ENERGETIC	LAZY
22. RELAXED	UPTIGHT
23. ALERT	UNALERT
24. GENTLE	HARD
25. COMFORTABLE	UNCOMFORTABLE
26. QUIET	LOUD
27. EASY GOING	STUBBORN
28. SLOW	FAST

The scoring procedure took about 1 hour per session. There were two sessions, one per stable. So the observers had to come back twice. There were 28 bipolar constructs on the scoring list (the full scoring list for unfamiliar observers in included in Appendix 20). The rating procedure was the same as in the first study (see section 7.1.3.1.2.4. for more detailed description).

8.1.3.3. Horse performance scale

The familiar observers were given a list of their horses and were asked to rank each individual horse on a nine point performance scale, 1 being poor performance and 9 being excellent performance (The scale is included in Appendix 19).

8.1.4. Data Analysis

Adapted RGT was used to collect the data from familiar and unfamiliar participants. Inter-rater agreement was calculated with Kendall's W Coefficient of Concordance. Kendall was also used when testing how consistent the raters were when rating the horses on separate constructs. We used ANOVA to determine the differences in the degree of agreement between different groups of observers. We also applied the t-test for various post-hoc calculations following ANOVA.

Multidimensional scaling (MDS) was used to calculate distances between the elements on a multidimensional perceptual map. In our study MDS produced two-dimensional models. We used linear regression with all the constructs and identified dimensions associating them with highly correlated constructs. As familiar and unfamiliar observers produced their own perceptual maps, we calculated the Spearman correlation coefficient for MDS scores to see whether the horses were positioned along the dimensions in a similar way by both observer groups. At the end we explored the relationship between the personality profiles of individual horses and performance ratings. SPSS 12.0, SigmaPlot 7 and Excel (Office XP) statistical packages were used to analyse the data and to produce figures and tables.

8.2. Results

8.2.1. Inter-observer agreement

8.2.1.1. Observers rating horses on all constructs

As in the first study we calculated Kendall's coefficient of concordance W to determine degree of agreement within the observer groups.

8.2.1.1.1. Stable 1

The degree of agreement in both observer groups was mostly significant ($p<0.05$) (Kendall's W with p values are included in Appendix 21). However Figure 8.1 shows a higher degree of agreement for familiar observers.

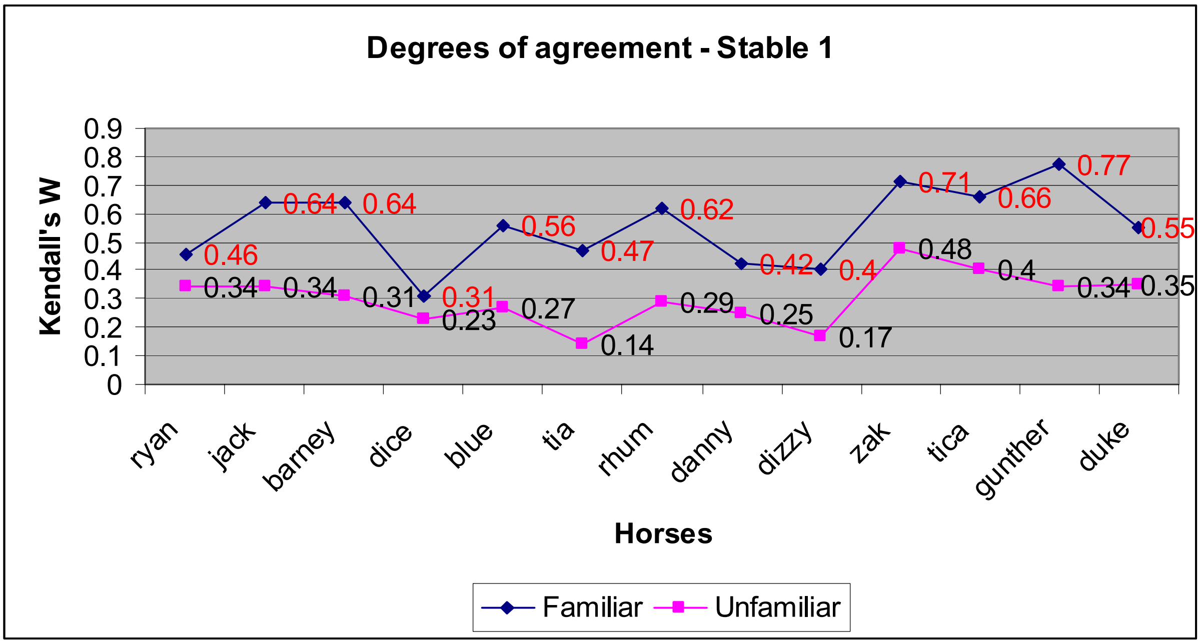


Figure 8.1

Kendall's W for Stable 1

The line graph (Figure 8.1) also shows that familiar observers reached a higher consensus for Zak and Gunther. Unfamiliar observers also agreed the most when they were rating Zak. For the lowest degrees of agreement the two observer groups chose different horses. Overall unfamiliar observers had a lower mean Kendall's W than familiar observers (Figure 8.2). Familiar group reached a mean W of 0.55 and the unfamiliar group a mean value W of 0.3.

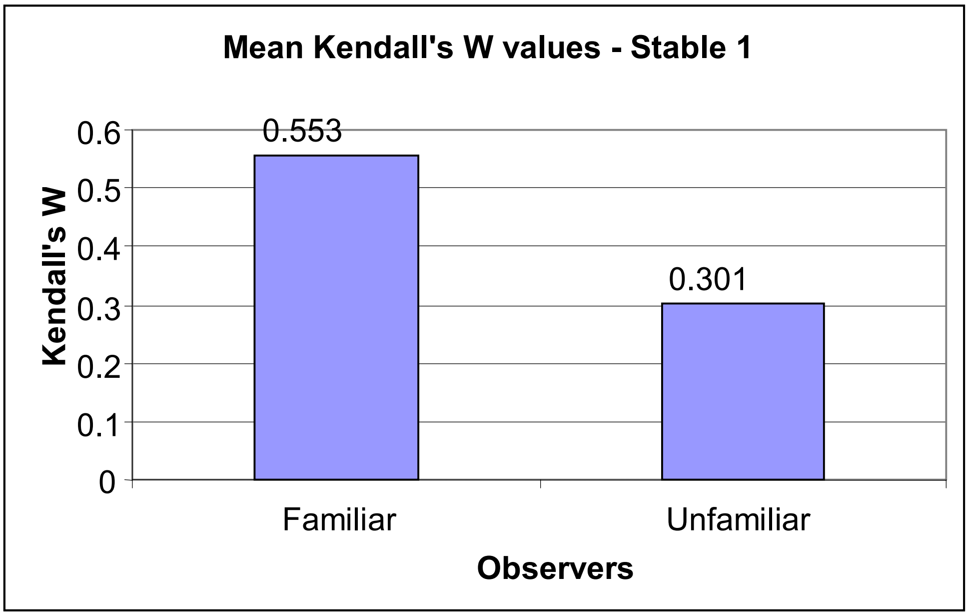


Figure 8.2
Mean Kendall's W for Stable 1

There was significant difference in the means between the two observer groups for Stable 1 ($t=9.536$, $P<0.001$).

8.2.1.1.2. Stable 2

Results for Stable 2 are less consistent. We can still see that familiar observers' Kendall's W is higher for all horses, except Kurt and Molly (Figure 8.3). Both observer groups seem have a lower degree of agreement for Lil and Candy, although the Kendall's W for Blossom is high within the familiar group and low within the unfamiliar group.

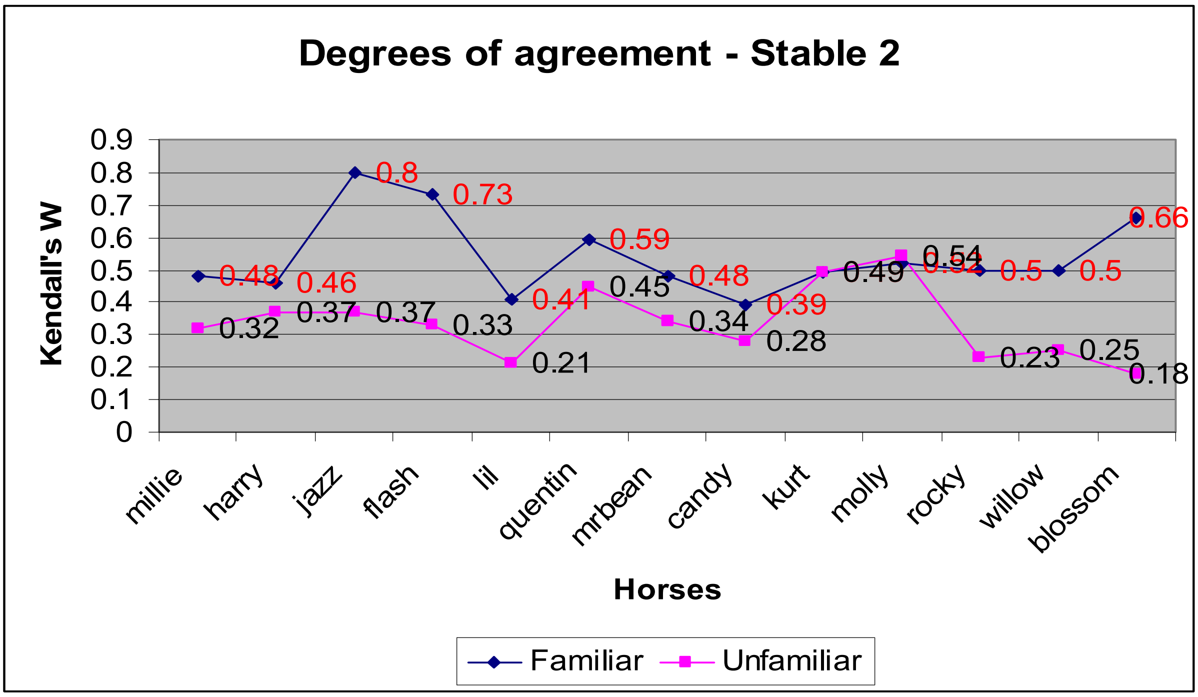


Figure 8.3

Kendall's W for Stable 2 – Familiar and unfamiliar observers

Overall Kendall's W for familiar observers in Stable 2 was 0.54 and for unfamiliar 0.33 (Figure 8.4).

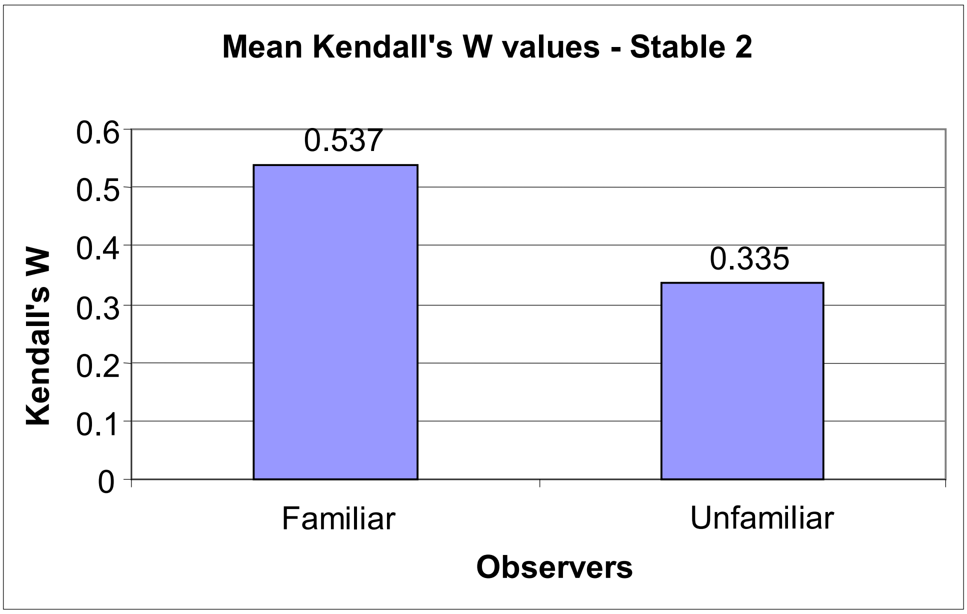


Figure 8.4
Mean Kendall's W for Stable 2 for both observer groups

There was significant difference in the means between the two observer groups for Stable 2 ($t=4.656$, $P<0.001$).

The results from the first two stables showed that familiar observers had a significantly higher Kendall's W than unfamiliar observers. The exact Kendall's W values are included in Appendix 21.

8.2.1.1.3. Stable 3

There was no video recorded at Stable 3. Familiar observers were the only group that rated their own horses. The degree of agreement was highly significant for all horses ($p<0.001$) (Appendix 21).

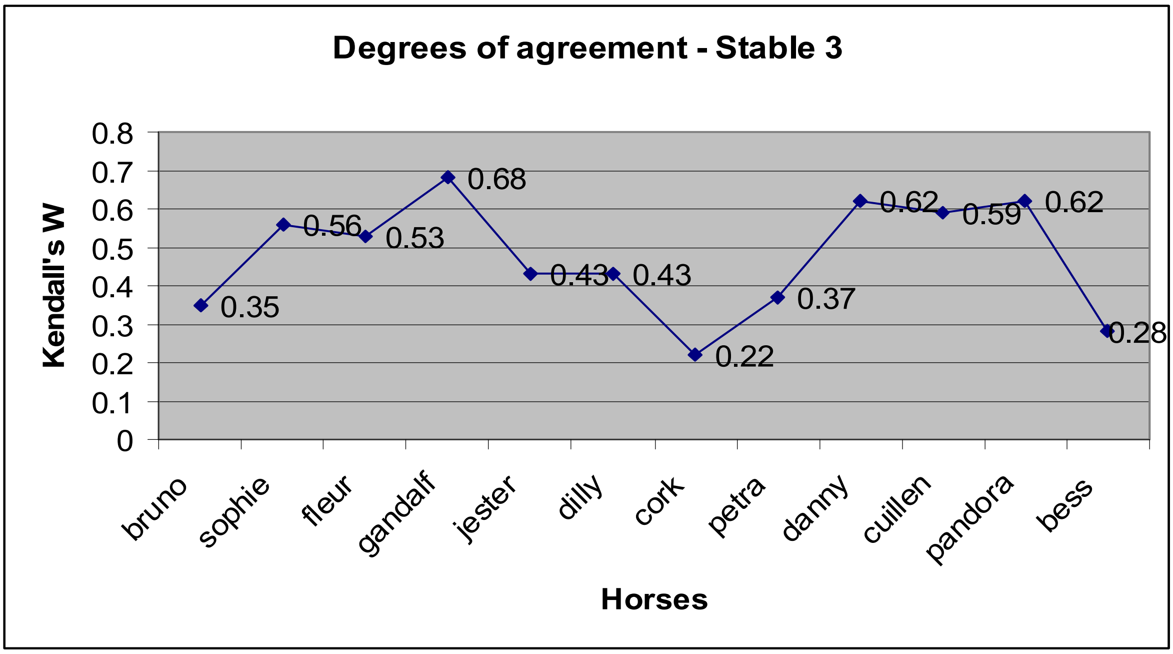


Figure 8.5
Kendall's W for Stable 3

The familiar observers from Stable 3 had the highest agreement for Gandalf, and the lowest for Cork. A high Kendall's W was also reached for Danny and Pandora, and a low one for Bruno and Bess (Figure 8.5). Most of the Kendall's W for other horses were between 0.4 and 0.6, which is similar to the scores of familiar observers from the first two stables.

8.2.1.2. Difference between elicited and partially provided constructs

The scoring lists for familiar and unfamiliar observer groups were divided into two parts. In the familiar group the first 16 constructs were provided by at least 12 observers, which is half of the number of observers that completed the study. The remaining 18 constructs were provided by at least seven observers. We were interested in whether there was any difference in the degree of agreement between the two groups of descriptors. In the unfamiliar observer group the first 14 constructs were elicited by at least eight participants, and the last 14 by seven participants or less.

8.2.1.2.1. Stable 1

The degree of agreement is significant for most of the horses, regardless of the mode of elicitation of the constructs (the complete list of Kendall's W with p-values is included in Appendix 22). There seems to be no consistent pattern for Kendall's W regarding elicited and provided constructs (Figure 8.6).

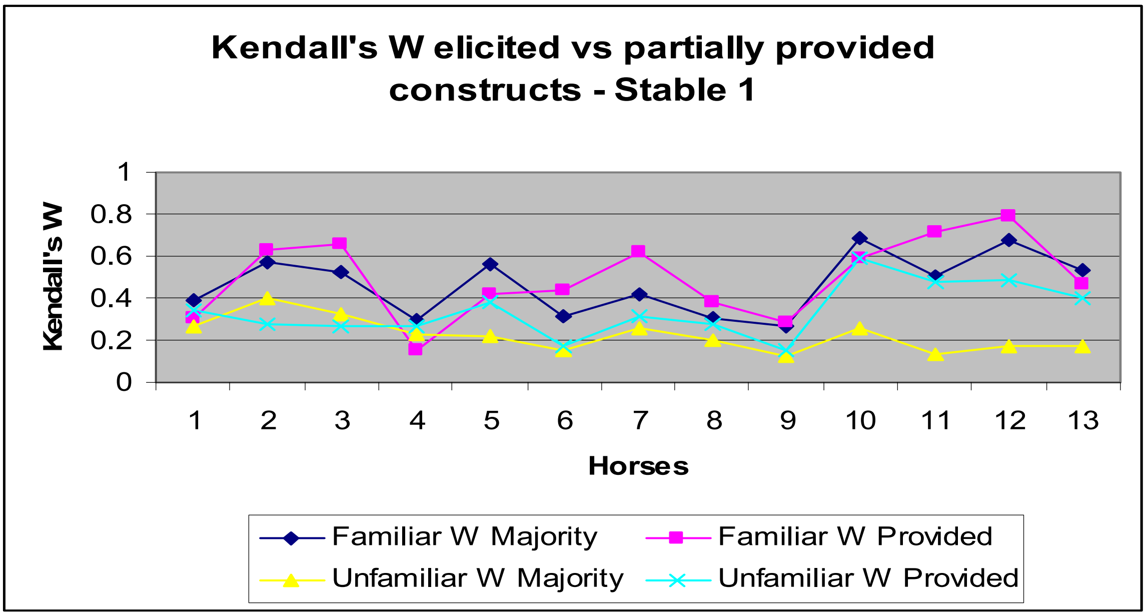


Figure 8.6

Kendall's W for elicited and partially provided constructs – Familiar and unfamiliar observers – Stable 1

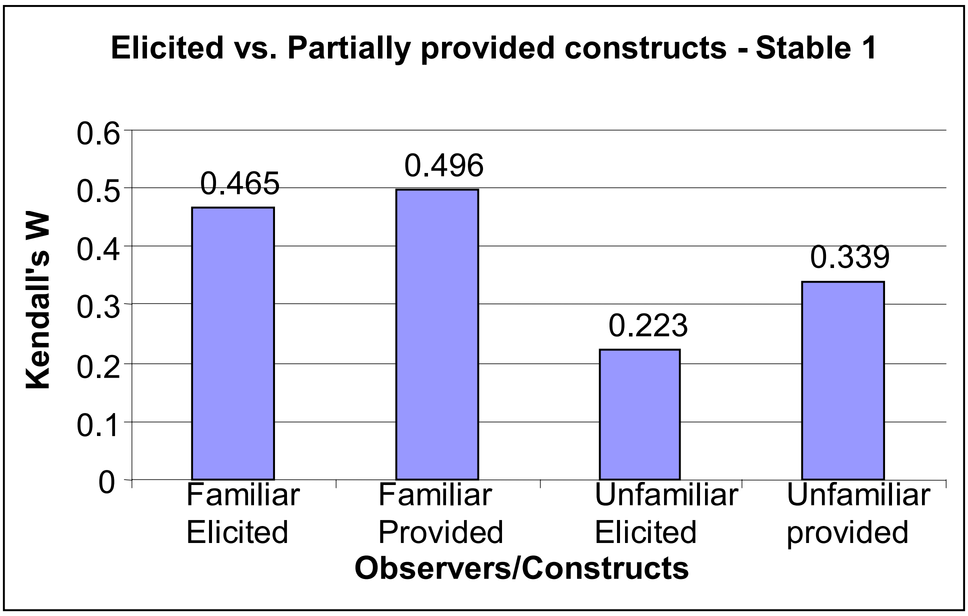


Figure 8.7

Mean Kendall's W for elicited and partially provided constructs – both observer groups – Stable 1

Comparing the mean values, the two observer groups agreed more when rating horses on provided constructs (Figure 8.7). The T-test showed that the agreement between familiar participants was not significantly influenced by elicitation ($t=0.896$, $p>0.05$). Unfamiliar observers agreed significantly more on provided constructs ($t=2.803$, $p<0.05$).

8.2.1.2.2. Stable 2

The degree of agreement is significant in most cases for the second stable as well (see Appendix 22). The line graph (Figure 8.8) again shows no consistent difference between provided and elicited constructs in any of the observer groups. For example, familiar observers had a higher degree of agreement on elicited constructs for horse one, but not for horses four, 12 and 13.

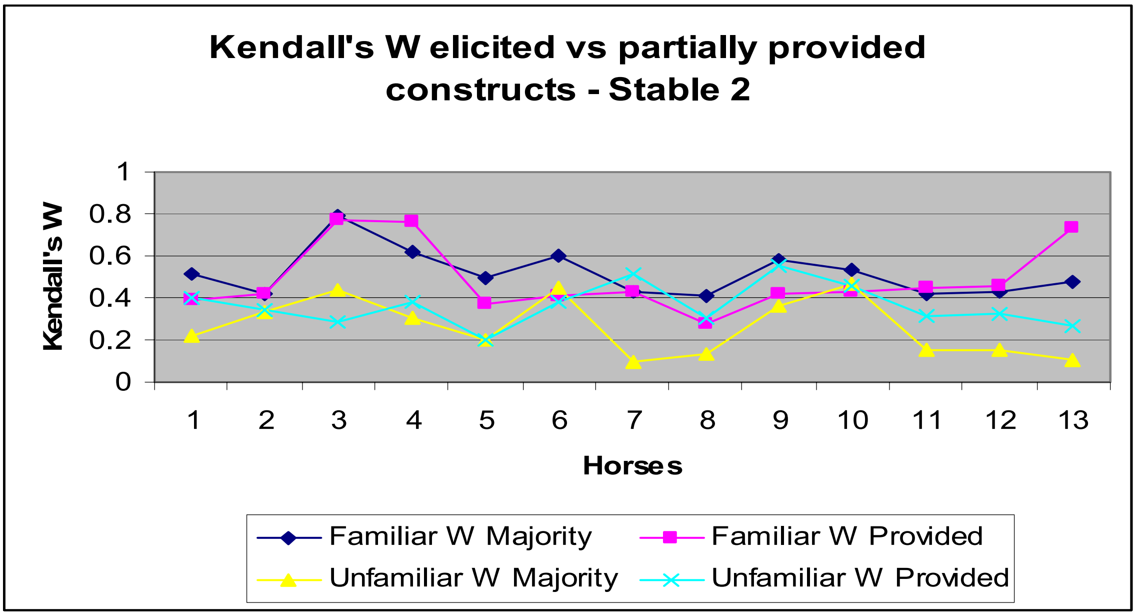


Figure 8.8

Kendall's W for elicited and partially provided constructs – Familiar and unfamiliar observers – Stable 2

The mean values indicate slightly higher Kendall's W coefficients when familiar observers scored the horses on constructs generated by the majority (Figure 8.9).

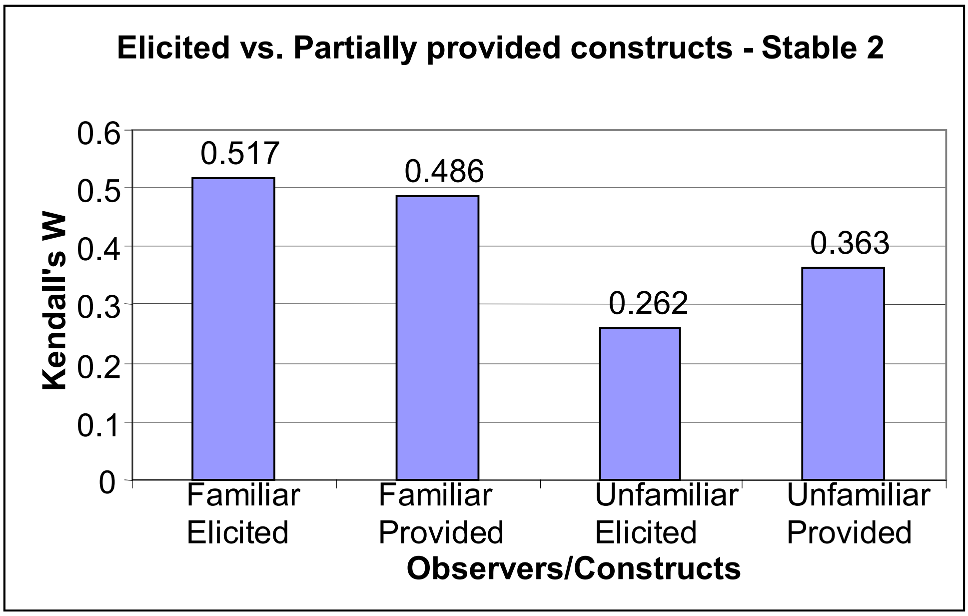


Figure 8.9
Mean Kendall's W for elicited and partially provided constructs – both observer groups – Stable 2

The difference between the two was however not significant (t=0.873, p>0.05). Unfamiliar observers again agreed significantly more on the provided constructs (t=2.507, p<0.05).

8.2.1.2.3. Stable 3

The horses from Stable 3 were rated only by the familiar raters. They showed a slightly higher degree of agreement when they were rating horses on constructs produced by the majority. There was no significant difference between the two groups of descriptors (t=0.424, p>0.05).

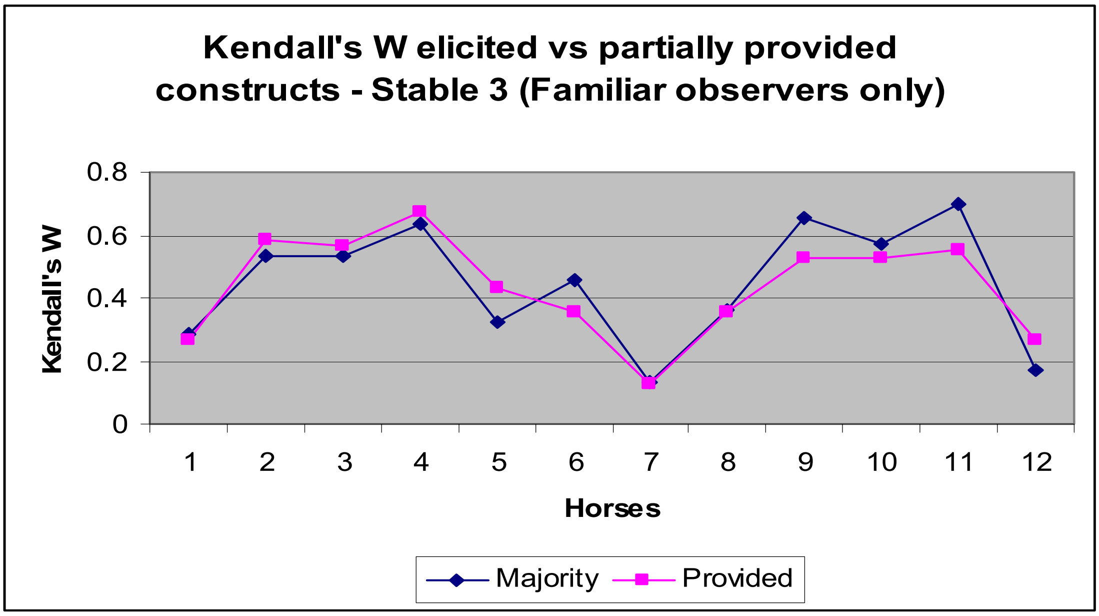


Figure 8.10

Kendall's W for elicited and partially provided constructs – Stable 3

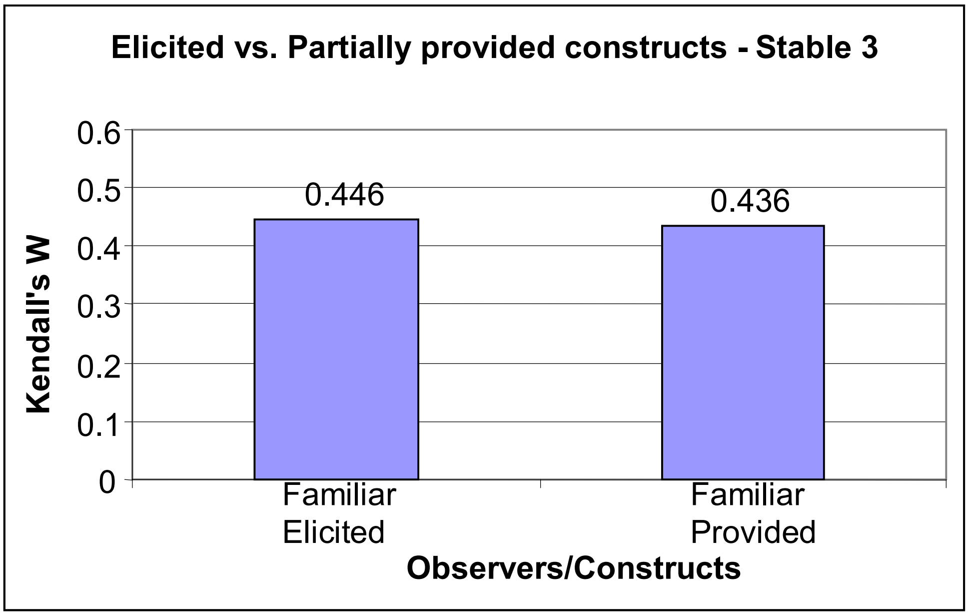


Figure 8.11

Mean Kendall's W for elicited and partially provided constructs – Stable 3

Looking at the results from three stables, it is not possible to identify consistency. It seems that familiar raters have a slightly higher agreement on constructs generated by the majority. Unfamiliar raters on the other hand agreed significantly more when the constructs were provided to them.

We were interested to investigate whether there were any differences in the degree of agreement between the first and the last three constructs on the scoring lists. These were constructs elicited by the higher and the lower number of observers respectively. First, there were no significant differences between these two groups of constructs, and secondly, there was no consistent pattern when comparing the mean Kendall's W. The full presentation of the results is included in Appendix 23.

8.2.1.3. Consistency of observers rating horses on separate constructs

We calculated the Kendall's W also for each separate construct. A higher degree of agreement between the users indicates consistent and systematic use. If the use of a construct was systematic and consistent we can conclude that it was understood by the observers and applied in a similar way when rating the horses. Kendall's W was calculated separately for each construct, for all observer groups. We ranked the ten most consistently used constructs per stable.

8.2.1.3.1. Familiar observers – all three stables

The familiar observers used the constructs consistently and systematically, which is indicated by the significant value of Kendall's W (Table 8.7). The Kendall's W for all constructs was significant. Therefore all familiar observers used the constructs consistently.

Table 8.7

Kendall's W for all constructs – Familiar observers – All 3 stables. Ten constructs with the highest Kendall's W are enumerated.

CONSTRUCTS		**Stable 1**	**Stable 2**	**Stable 3**
1. FRIENDLY	UNFRIENDLY	.422 ***	.377 **	.483***
2. ATTENTION SEEKER	INDEPENDENT	.317 **	.327 **	.470***
3. CONFIDENT	TIMID	.360***	.457***	**.630***1**
4. BOLD	SHY	.405***	.588*****4**	.592*****2**
5. EASY TO HANDLE	DIFFICULT	.626*****2**	.445***	.490*****10**
6. SOCIABLE	UNSOCIABLE	.432***	.500***	.420***
7. BOSSY	SUBMISSIVE	.391***	.514*****9**	.558*****4**

8. HAPPY	UNHAPPY	.402***	.564*****6**	.400***
9. LAID BACK	HIGHLY STRUNG	.451***	.489***	.343***
10. CALM	EXCITABLE	.459***	.499***	.350***
11. PLAYFUL	BORING	.282 **	.663*****2**	.497*****8**
12. AFFECTIONATE	AGGRESSIVE	.354***	.502*****10**	.377***
13. OBEDIENT	DISOBEDIENT	.268 **	.238 *	.303 **
14. NERVOUS	BOLD	.273 **	.538*****7**	.527*****5**
15. INTERESTED	DISINTERESTED	.322 **	.423***	.298 **
16. AGGRESSIVE	GENTLE	.486***	.332***	.375***
17. DOMINANT	SUBMISSIVE	.523***10	.412***	.376***
18. GENTLE	ROUGH	.594*****3**	.348***	.506***7
19. EXPERIENCED	INEXPERIENCED	.552*****5**	.621*****3**	.465***
20. MATURE	IMMATURE	.551*****6**	.699*****1**	.567*****3**
21. EASY TO WORK WITH	DIFFICULT	.547*****7**	.410***	.466***
22. CHEEKY	WELL BEHAVED	.512***	.368***	.495***9
23. SECURE	INSECURE	.378 **	.374***	.473***
24. NASTY	NICE	.522***	.426***	.380***
25. BRAVE	SCARED	.351***	.487***	.369***
26. FORWARD	HESITANT	.413***	.426***	.287 **
27. BULLY	BULLIED	.537*****9**	.334***	.558*****5**
28. STUBBORN	WILLING	.275 **	.224 *	.207 *
29. INTELLIGENT	THICK	.338*****8**	.303*	.263 **
30. RELAXED	TENSE	.557*****4**	.579*****5**	.430***
31. CONTENT	GRUMPY	.417***	.434***	.375***

32. PATIENT	IMPATIENT	.515***	.336***	.467***
33. CLEVER	STUPID	.323***	.244 *	.309***
34. BARGY	QUIET	.677*****1**	.267 *	.546*****6**
35. PERFORMANCE		.242 *	.523*****8**	.375***

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

The line graph below (Figure 8.12) shows how consistently and systematically the observers used the constructs. There is no apparent pattern, although some constructs (e.g. construct 20 "mature-immature") seem to be used more systematically by all observer groups than others (e.g. construct 28 "stubborn-willing").

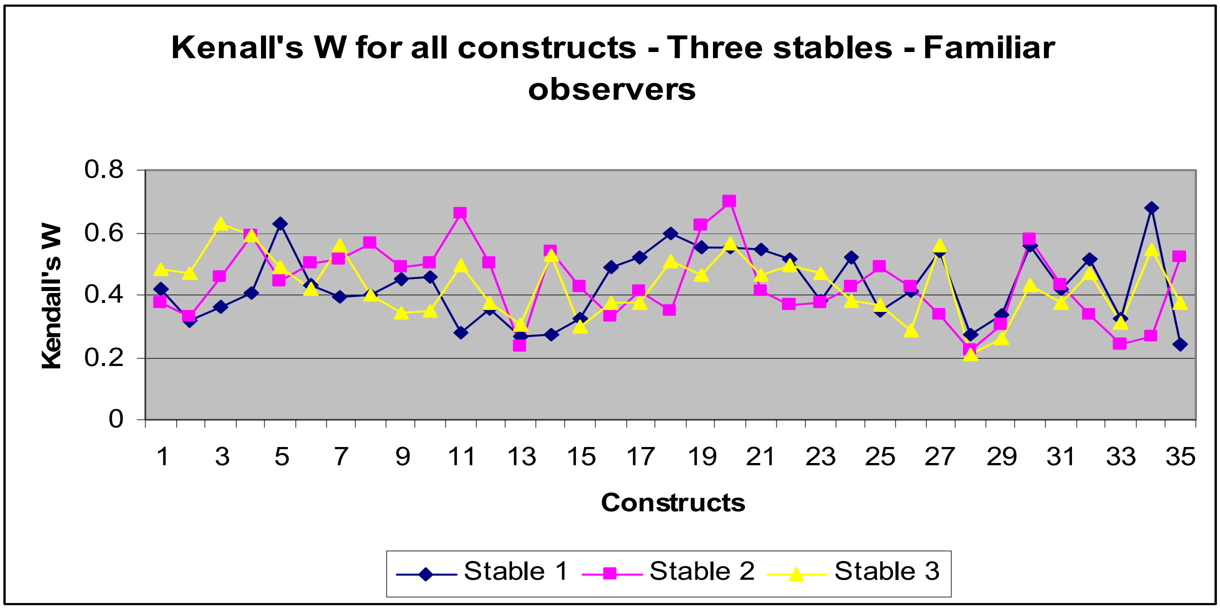


Figure 8.12

Mean Kendall's W for all constructs – Familiar observers – All 3 stables

To see which constructs were used systematically by the observers from all three stables, we rank ordered the constructs according to their Kendall's W value (Table 8.8).

Table 8.8

Ordering of 10 constructs with the highest Kendall's W – Familiar observers – All 3 stables

CONSTRUCTS		**Stable 1**	**Stable 2**	**Stable 3**
BARGY	QUIET	**1**		**6**
EASY TO HANDLE	DIFFICULT	**2**		**10**
GENTLE	ROUGH	**3**		**7**
RELAXED	TENSE	**4**	**5**	
EXPERIENCED	INEXPERIENCED	**5**	**3**	
MATURE	IMMATURE	**6**	**1**	**3**
EASY TO WORK WITH	DIFFICULT	**7**		
INTELLIGENT	THICK	**8**		
BULLY	BULLIED	**9**		**5**
DOMINANT	SUBMISSIVE	**10**		
CONFIDENT	TIMID			**1**
PLAYFUL	BORING		**2**	**8**
BOLD	SHY		**4**	**2**
HAPPY	UNHAPPY		**6**	
NERVOUS	BOLD		**7**	**5**
PERFORMANCE			**8**	
BOSSY	SUBMISSIVE		**9**	**4**
AFFECTIONATE	AGGRESSIVE		**10**	
CHEEKY	WELL BEHAVED			**9**

While some constructs, like "mature-immature" were used consistently by all three groups of familiar observers, the others, for example "affectionate-aggressive" or "cheeky-well behaved" were used more systematically only by one group of observers.

Three groups of familiar observers, one at each stable, generated the uniform scoring list. We were interested in seeing whether there was a correlation of Kendall's W between the stables. A significant correlation would indicate consistent use of all constructs across the three groups of familiar observers. The Pearson correlations coefficient shows that there is a significant correlation between Stable 1 and Stable 3 ($p<0.05$) and Stables 2 and 3 ($p<0.05$), but not between Stables 1 and 2 (Table 8.9).

Table 8.9

Pearson correlation for Kendall's W for all constructs – Familiar observers

		Stable2	Stable3
Pearson Correlation	**Stable 1**	**.027 NS**	**.382***
	Stable 2	**1**	**.366***

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

8.2.1.3.2. Unfamiliar observers

We also explored the consistency of construct use by the unfamiliar observers. In contrast to familiar observers, the unfamiliar observers' Kendall's W was not significant for some constructs, e.g. "calm-nervous" and "laid back-worried" (Table 8.10).

Table 8.10

Kendall's W for all constructs – Unfamiliar observers – Stables 1 and 2

CONSTRUCTS		**Stable 1**	**Stable 2**
1. INTERESTED	UNINTERESTED	.237 **	.554***
2. BOLD	SHY	.420***	.473***
3. CURIOUS	INDIFFERENT	.282***	.438***
4. PLAYFUL	BORED	.382***	.455***
5. OUTGOING	**TIMID**	**.351*****	**.530*****
6. CALM	NERVOUS	.140 NS	.147 NS
7. HAPPY	SAD	.315***	.349***
8. ENTERTAINED	BORED	.305***	.416***
9. AT EASE	SCARED	.096 NS	.169 NS
10. INQUISITIVE	DISINTERESTED	.370***	.519***
11. CONFIDENT	WARY	.366***	.406***
12. FRIENDLY	UNFRIENDLY	.307***	.492***
13. BRAVE	SCARED	.243**	.313***
14. LAID-BACK	WORRIED	.097 NS	.171 NS
15. CALM	BOISTEROUS	.284***	.373***
16. AFFECTIONATE	COLD	.270***	.556***
17. STEADY	JUMPY	.185 *	.247 *
18. NOSY	NOT BOTHERED	.321***	.472***
19. STRONG	ANXIOUS	.243 **	.323***
20. DOMINANT	SUBMISSIVE	.373***	.357***
21. ENERGETIC	LAZY	.463***	.546***
22. RELAXED	UPTIGHT	.144 NS	.229 **
23. ALERT	UNALERT	.280***	.285***
24. GENTLE	HARD	.324***	.349***
25. COMFORTABLE	UNCOMFORTABLE	.127 NS	.228 *
26. QUIET	LOUD	.370***	.491***
27. EASY GOING	STUBBORN	.218 **	.246 **
28. SLOW	FAST	.395***	.381***

*** P<0.001;

**P<0.01; *P<0.05; NS – not significant

The constructs that were used most systematically by unfamiliar observers were "bold-shy", "energetic-lazy". The constructs used least systematically were "calm-nervous" and "at ease-scared" (Table 8.11).

Table 8.11
Ordering of 10 constructs with the highest Kendall's W – Unfamiliar observers

CONSTRUCTS		**Stable 1**	**Stable 2**
ENERGETIC	LAZY	1	3
BOLD	SHY	2	8
SLOW	FAST	3	
PLAYFUL	BORED	4	10
DOMINANT	SUBMISSIVE	5	
QUIET	LOUD	6	7
INQUISITIVE	DISINTERESTED	7	5
CONFIDENT	WARY	8	
OUTGOING	TIMID	9	4
GENTLE	HARD	10	
AFFECTIONATE	COLD		1
INTERESTED	UNINTERESTED		2
FRIENDLY	UNFRIENDLY		6
NOSY	NOT BOTHERED		9

Constructs nine ("at ease-scared") and 14 ("laid back-worried") were used less consistently when rating horses from both stables (Figure 8.13).

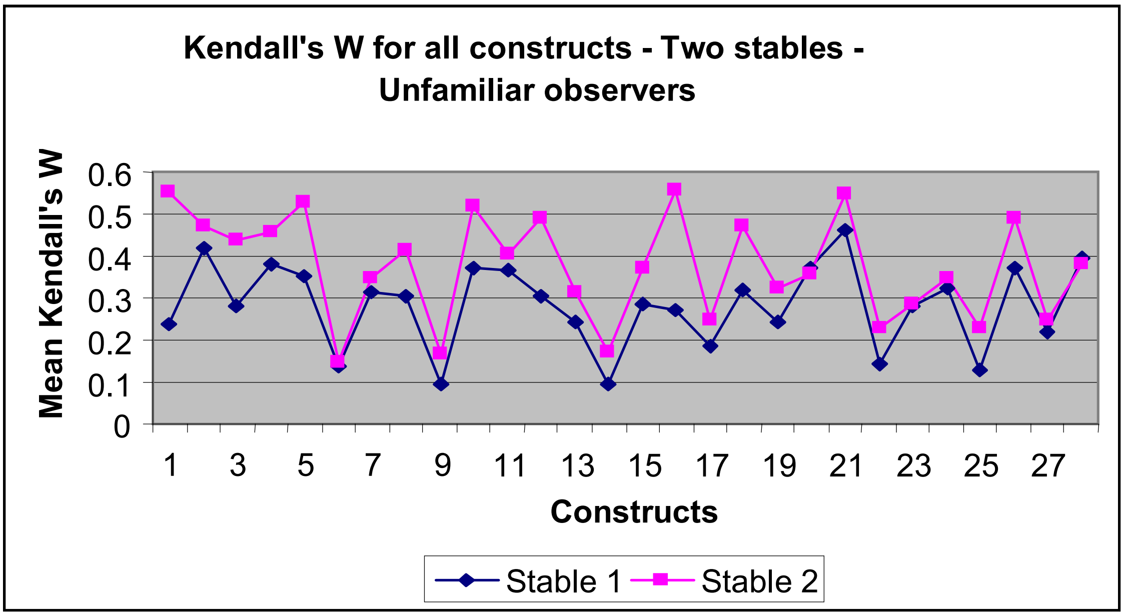


Figure 8.13

Mean Kendall's W for all constructs – Unfamiliar observers – Stables 1 and 2

The Kendall's W was significantly higher when observers used constructs to rate horses from Stable 2 ($t=6.153$; $p<0.001$) (Figure 8.13). The unfamiliar observers therefore used the constructs more consistently when they were observing horses from Stable 2 (Figure 8.14).

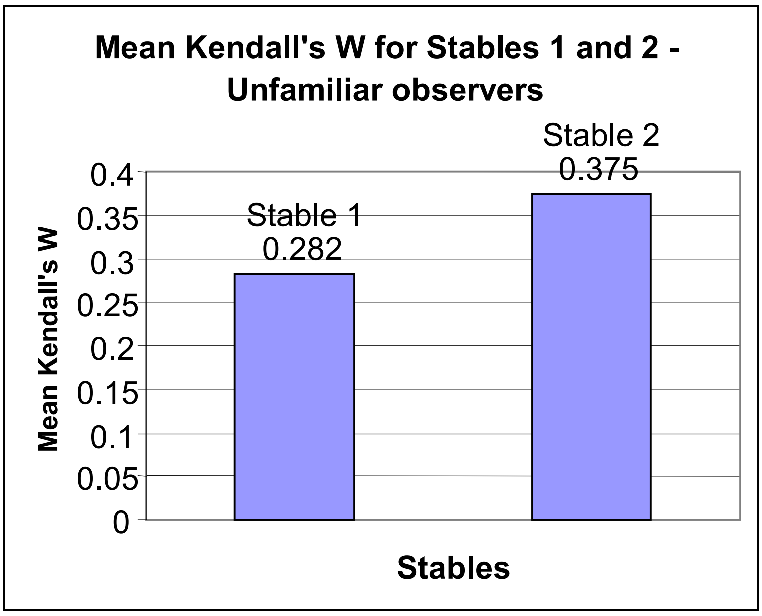


Figure 8.14
Mean Kendall's W for all constructs – Unfamiliar observers – Stables 1 and 2

Even though there were differences in the two means, the correlation of scores was still significant (p<0.01). That means that dynamic of applying different constructs was similar (Table 8.12).

Table 8.12

Pearson correlation for Kendall's W for all constructs – Unfamiliar observers

Pearson Correlation		**Stable 2**
	Stable 1	**.771****

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

8.2.2. Discovering and identifying horse personality dimensions

8.2.2.1. Stable 1

To uncover personality dimensions in horses, Multidimensional scaling (MDS) was conducted on both sets of data. MDS is used to calculate distances between objects in a spatial map. Objects can be animals, concepts or people. MDS is further described in section 7.2.3. The main output of MDS is spatial configuration of points or horses in two or more dimensional space. To obtain valid results it is recommended to have at least four elements per dimension, nine per two dimensions, 13 per three and so on. To identify the dimensions we associated the constructs with two dimensions using linear regression. At the end we correlated the dimensions produced by familiar and unfamiliar observers.

8.2.2.1.1. Multidimensional scaling (MDS)

Both the familiar and unfamiliar observers positioned the horses along the two dimensions by rating them on personality adjectives they partially generated themselves (Figures 8.15 and 8.16).

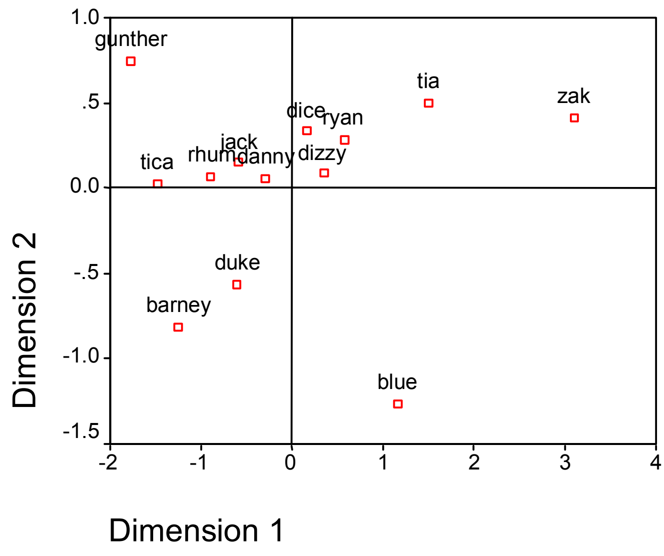


Figure 8.15

Euclidean two-dimensional distance model produced by familiar observers

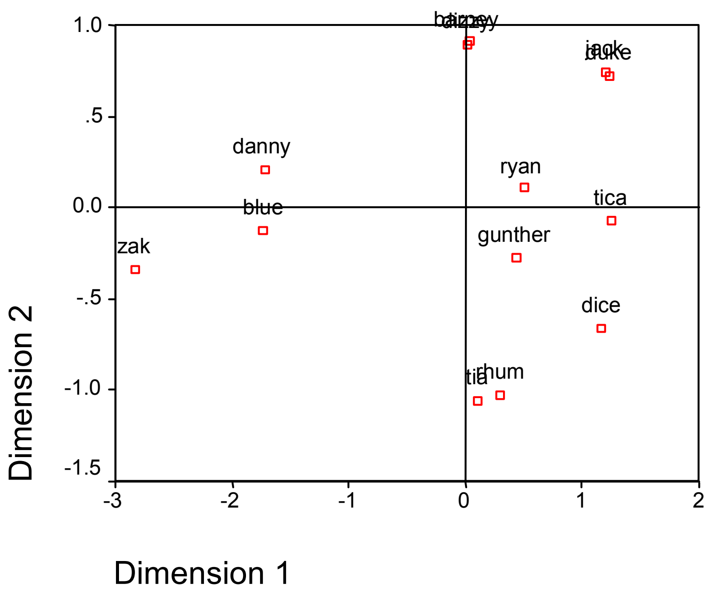


Figure 8.16

Euclidean two-dimensional distance model produced by unfamiliar observers

To look at the proportion of variance of the scaled data (disparities) which is accounted for by their corresponding distances on the matrix we calculated Squared correlation in distances (RSQ). The RSQ for the familiar observers in Stable 1 was 0.97564, and for the unfamiliar observers 0.94103. The high value corresponds to the high proportion of variance explained. We were also interested in seeing how well the model fits. The goodness of fit measure is indicated by stress function. For that purpose we calculated stress values using Kruskal's stress formula 1, which is expressed in relation to the size of the real values X. Stress for the matrix produced by the familiar observers was 0.08085. Stress for unfamiliar observers was 0.11750. The values range from 0-1. The lower the value the better the fit.

8.2.2.1.2. Identification of dimensions

To identify the dimensions in both MDS perceptual maps, we conducted linear regression to associate mean scores for each construct with the dimensions.

Data from Stable 1 generated by the familiar group indicates that the first dimension is associated with constructs that form the *Compliance facet* of the agreeableness dimension (Costa & McCrae, 1992). These constructs are "easy to work with-difficult", "easy to handle-difficult" "cheeky-well behaved", "patient-impatient". Highly associated with the first dimension were also the constructs classified under the "tender-mindedness" and "altruism" facets of A: e.g. "aggressive-gentle", "gentle-rough" and "nasty-nice". Therefore the first dimension was labelled *Agreeableness* (A). The second dimension is not very clearly defined. The constructs "interested-disinterested", "calm-excitable" and "sociable-unsociable" that are significantly associated with dimension 2 most likely represent the *warmth* facet of the extroversion dimension, so we labelled the second dimension *Extroversion* (E). Dimension 1 is also associated with constructs that could be classified either under neuroticism (N), e.g. "laid back-highly strung" or E, e.g. "friendly-unfriendly" and "bargy-quiet" (Table 8.13).

Table 8.13
Linear regression to identify two dimensions – Stable 1 – Familiar observers

CONSTRUCTS		T – Regression coefficient	
		Stable 1	
		DIM 1	DIM 2
FRIENDLY	UNFRIENDLY	**6.298*** **	2.484*
ATTENTION SEEKER	INDEPENDENT	-1.192 NS	**2.769***
CONFIDENT	TIMID	-1.921 NS	.184 NS
BOLD	SHY	-2.920*	.169 NS
EASY TO HANDLE	DIFFICULT	**13.354*** **	1.211 NS
SOCIABLE	UNSOCIABLE	**7.198*** **	**3.794** **
BOSSY	SUBMISSIVE	-3.778**	-.229 NS
HAPPY	UNHAPPY	2.308*	2.106 NS
LAID BACK	HIGHLY STRUNG	**7.335*** **	-2.772*
CALM	EXCITABLE	**6.445*** **	**-5.147*** **
PLAYFUL	BORING	-3.498**	3.812**
AFFECTIONATE	AGGRESSIVE	4.138**	2.858*

OBEDIENT	DISOBEDIENT	**5.780*** **	-.078 NS
NERVOUS	BOLD	1.037NS	1.271 NS
INTERESTED	DISINTERESTED	**-5.533*** **	**5.268*** **
AGGRESSIVE	GENTLE	**-10.528*** **	**-3.457** **
DOMINANT	SUBMISSIVE	**-5.000*** **	-.619 NS
GENTLE	ROUGH	**12.818*** **	**3.154** **
EXPERIENCED	INEXPERIENCED	1.320 NS	.760 NS
MATURE	IMMATURE	1.838NS	-.216 NS
EASY TO WORK WITH	DIFFICULT	**14.564*** **	.644 NS
CHEEKY	WELL BEHAVED	**-8.174*** **	.287 NS
SECURE	INSECURE	1.824NS	-.029 NS
NASTY	NICE	**-8.556*** **	-3.111*
BRAVE	SCARED	-1.331NS	-.410 NS
FORWARD	HESITANT	**-5.035*** **	.716 NS
BULLY	BULLIED	-3.482**	-1.332 NS
STUBBORN	WILLING	-1.567NS	-1.839 NS
INTELLIGENT	THICK	-2.697*	2.630*
RELAXED	TENSE	**4.742*** **	-3.129*
CONTENT	GRUMPY	**6.342*** **	.064 NS
PATIENT	IMPATIENT	**8.689*** **	**-3.863** **
CLEVER	STUPID	-2.414*	3.065*
BARGY	QUIET	**-10.382*** **	1.002 NS

P<0.05; ** P<0.01; *** P<0.001; NS - not significant.

Linear regression for the unfamiliar MDS model indicates more clearly defined dimensions. The constructs highly associated with Dimension 1 are "bold-shy", "confident-wary", "energetic-lazy", "slow-fast", "happy-sad" and "outgoing-timid". These adjectives could be associated with either with extroversion or neuroticism. We labelled the first dimension *Neuroticism* (N), because constructs like "brave-scared" and "calm-boisterous" were also significantly associated with this dimension. The second dimension was labelled *Extroversion* (E) and was significantly associated with the constructs "affectionate-cold", "friendly-unfriendly", and "interested-uninterested" (Table 8.14). There is clearly an overlap between the dimensions, so the labelling is arbitrary.

Table 8.14

Linear regression for the unfamiliar observers – Stable 1

CONSTRUCTS		T – Regression coefficient	
		Stable 1	
		DIM 1	DIM 2
INTERESTED	UNINTERESTED	2.401*	**6.574*** **
BOLD	SHY	**7.688*** **	2.097 NS
CURIOUS	INDIFFERENT	1.575 NS	**4.437*** **
PLAYFUL	BORED	**5.355*** **	2.300*
OUTGOING	**TIMID**	**7.265*** **	.952 NS
CALM	NERVOUS	2.674*	-.154 NS
HAPPY	SAD	**5.750*** **	1.741 NS
ENTERTAINED	BORED	3.385**	2.072 NS
AT EASE	SCARED	3.490**	.775 NS
INQUISITIVE	DISINTERESTED	3.003*	**3.299** **
CONFIDENT	WARY	**8.094*** **	1.480 NS
FRIENDLY	UNFRIENDLY	.696 NS	**4.299** **
BRAVE	SCARED	**4.965*** **	.803 NS
LAID-BACK	WORRIED	1.740 NS	-.049 NS
CALM	BOISTEROUS	**-4.918*** **	.482 NS
AFFECTIONATE	COLD	.279 NS	**5.101*** **
STEADY	JUMPY	.558 NS	.608 NS
NOSY	NOT BOTHERED	3.793**	2.547*
STRONG	ANXIOUS	**4.277** **	-.840 NS
DOMINANT	SUBMISSIVE	**5.721*** **	-.164 NS
ENERGETIC	LAZY	**9.028*** **	-.942 NS
RELAXED	UPTIGHT	3.434**	.836 NS
ALERT	UNALERT	3.650**	1.442 NS
GENTLE	HARD	-1.886 NS	**2.342** **
COMFORTABLE	UNCOMFORTABLE	4.042**	-.502 NS
QUIET	LOUD	**-7.417*** **	2.076 NS
EASY GOING	STUBBORN	.295 NS	1.779 NS
SLOW	FAST	**-16.001*** **	**4.908*** **

P<0.05; ** P<0.01; *** P<0.001; NS - not significant.

The two dimensions produced by familiar observers were labelled A and E. The dimensions generated by unfamiliar observers were labelled N and E. The dimensions overlapped and the adjectives could not be clearly classified. The next step was to cross-validate the dimensions by correlating horse personality scores across the two observer groups.

8.2.2.1.3. Correlation of dimensions

The correlation of horse personality scores generated by both observer groups is significant for Dimension 1 (rho=0.55, $p<0.05$). It has been labelled A in the familiar group and N in the unfamiliar group. There was no significant correlation for the second dimension, labelled E in both observer groups. However, the correlation coefficient indicates a tendency towards significance (Table 8.15).

Table 8.15

Spearman correlation coefficient for Stable 1

Spearman rho	**Dimension 1 Unfamiliar**	**Dimension 2 Unfamiliar**
Dimension 1 Familiar	.55*	.275 NS
Dimension 2 Familiar	.022 NS	.522 NS p = .067

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

8.2.2.2. Stable 2

8.2.2.2.1. MDS

Horse personality scores generated by both the familiar and the unfamiliar observers also resulted in a spatial configuration of horses (Figures 8.17 and 8.18). Looking at the perceptual maps, there are clear differences in the way familiar and unfamiliar observers positioned the horses along the two dimensions. For example familiar observers located Kurt on the negative pole of dimension 1 and quite close to Lil and Harry (Figure 8.17). Unfamiliar observers located the same horse on the far positive end of dimension 1 and quite far away from Lil and Harry (Figure 8.18).

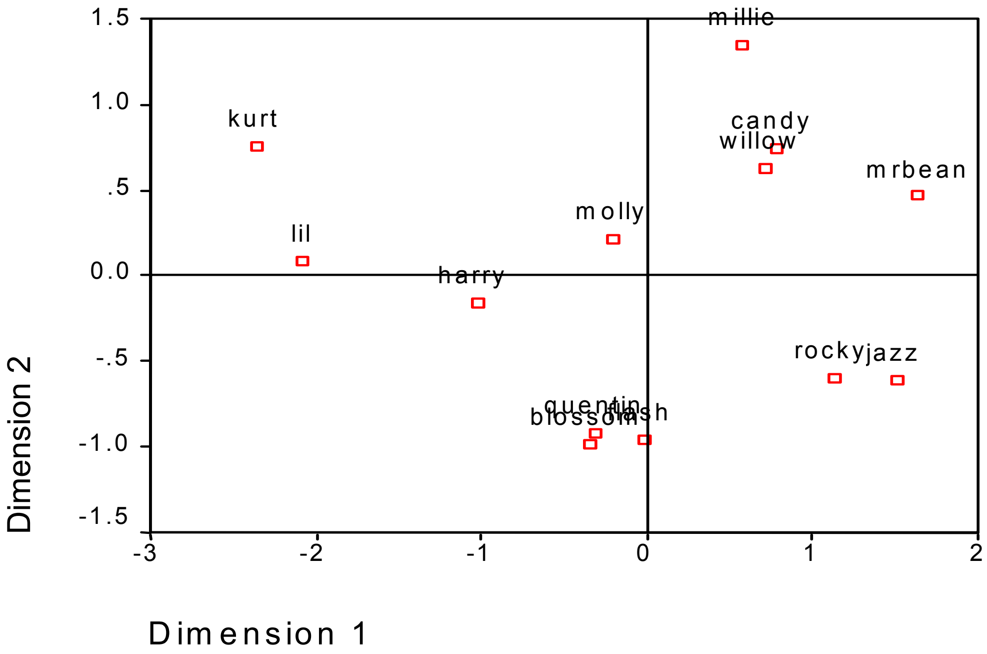


Figure 8.17

Euclidean two-dimensional distance model produced by familiar observers

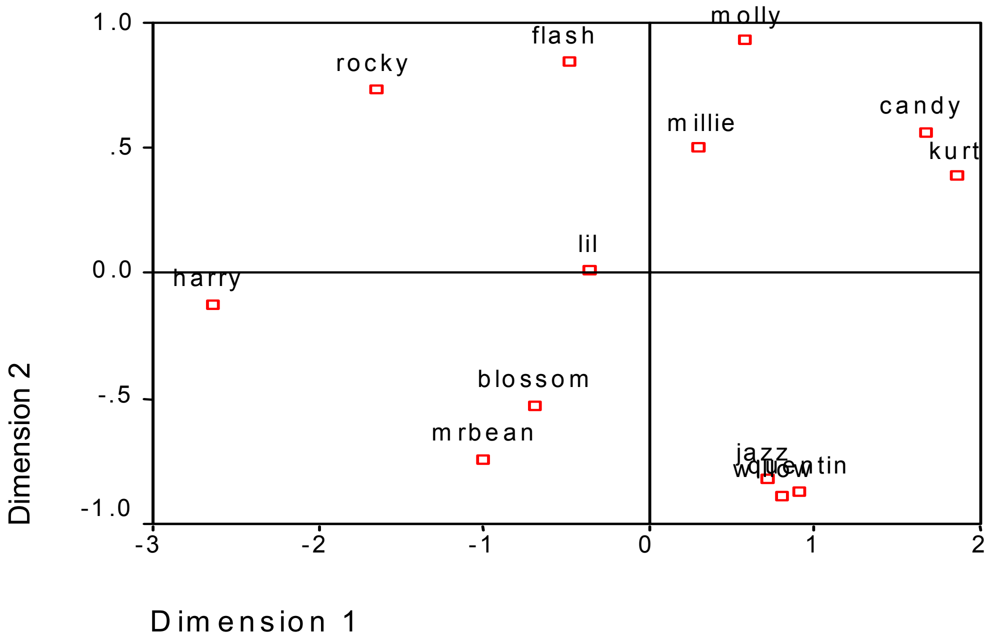


Figure 8.18

Euclidean two-dimensional distance model produced by unfamiliar observers

For familiar observers the RSQ of 0.90581 shows that most of the data is accounted for. The model also indicates low error measurement (stress=0.13449). Similarly for unfamiliar observers the RSQ = 0.92630 and the goodness of fit is indicated by the stress value of 0.12007.

8.2.2.2.2. Identification of dimensions

Linear regression on all constructs and two dimensional MDS models was calculated also on personality scores for horses from Stable 2. In the familiar group the constructs significantly associated with the first dimension were "nervous-bold", "brave-scared", "secure-insecure", "bold-shy", "confident-timid" and also "easy to handle-difficult". Apart from the last construct all others make up the anxiety and vulnerability facets of neuroticism dimension, therefore we labelled dimension 1 *Neuroticism* (N). The second dimension was associated with "affectionate-aggressive", "sociable-unsociable", "interested-disinterested", "nasty-nice" and "gentle-rough" (Table 8.15). These constructs make up the warmth and gregariousness facets of the extroversion dimension according to the Five Factor Model of personality (Costa & McCrae, 1992). Therefore we labelled the second dimension produced by familiar observers *Extroversion* (E).

Table 8.15

Linear regression to identify two dimensions - Stable 2 – Familiar observers

CONSTRUCTS		T – Regression coefficient	
		Stable 2	
		DIM 1	DIM 2
FRIENDLY	UNFRIENDLY	-1.040 NS	2.812*
ATTENTION SEEKER	INDEPENDENT	-.987 NS	**3.143****
CONFIDENT	TIMID	**-7.766*****	-2.558*
BOLD	SHY	**-8.539*****	-2.924*
EASY TO HANDLE	DIFFICULT	**-5.685*****	1.266 NS
SOCIABLE	UNSOCIABLE	-2.691*	**4.563*****
BOSSY	SUBMISSIVE	-3.359**	-1.303 NS
HAPPY	UNHAPPY	-2.511*	4.140**
LAID BACK	HIGHLY STRUNG	-2.682*	.791 NS
CALM	EXCITABLE	-2.110 NS	.560 NS
PLAYFUL	BORING	-1.078 NS	**2.585***
AFFECTIONATE	AGGRESSIVE	.042 NS	**4.837*****
OBEDIENT	DISOBEDIENT	-2.835*	2.348*
NERVOUS	BOLD	**8.047*****	3.190**
INTERESTED	DISINTERESTED	-1.334 NS	**4.595*****
AGGRESSIVE	GENTLE	.270 NS	**-3.614****
DOMINANT	SUBMISSIVE	-2.440*	-2.017 NS
GENTLE	ROUGH	.333 NS	**4.144****
EXPERIENCED	INEXPERIENCED	**-7.674*****	-3.803**
MATURE	IMMATURE	**-5.619*****	-3.047*
EASY TO WORK WITH	DIFFICULT	-3.872**	1.845 NS
CHEEKY	WELL BEHAVED	.106 NS	.488 NS
SECURE	INSECURE	**-12.126*****	.926 NS
NASTY	NICE	-.067 NS	**-4.565*****

BRAVE	SCARED	**-8.922*** **	-.324 NS
FORWARD	HESITANT	**-6.971*** **	.165 NS
BULLY	BULLIED	-2.559*	.218 NS
STUBBORN	WILLING	1.017 NS	**-4.197** **
INTELLIGENT	THICK	-2.716*	-.290 NS
RELAXED	TENSE	**-4.983*** **	.282 NS
CONTENT	GRUMPY	-3.096**	**5.720*** **
PATIENT	IMPATIENT	-1.697 NS	1.381 NS
CLEVER	STUPID	**-4.792*** **	.313 NS
BARGY	QUIET	-.025 NS	-1.557 NS

P<0.05; ** P<0.01; *** P<0.001; NS - not significant.

In the group of unfamiliar observers the two dimensions were again not clearly separated. The first dimension was significantly associated with "friendly-unfriendly", "affectionate-cold", "playful-bored", "interested-disinterested" and "bold-shy". These adjectives make up the *assertiveness* and *warmth* facets of *extroversion*. Therefore the first dimension has been labelled *Extroversion (E)*. The second dimension, however, was not clearly differentiated from the first. The constructs associated with it were "gentle-hard", "dominant-submissive", "affectionate-cold" and "friendly-unfriendly" (Table 8.16). They could be classified either as extroversion or neuroticism adjectives. We labelled the second dimension generated by unfamiliar observers the *Extroversion/Neuroticism* dimension (E/N).

Table 8.16

Linear regression to identify two dimensions - Stable 2 – Unfamiliar observers

CONSTRUCTS		T – Regression coefficient	
		Stable 2	
		DIM 1	DIM 2
INTERESTED	UNINTERESTED	**10.923***	-2.230*
BOLD	SHY	**10.651***	3.596**
CURIOUS	INDIFFERENT	**7.430***	-3.502**
PLAYFUL	BORED	**11.297***	-.367 NS
OUTGOING	**TIMID**	**9.202***	2.956*
CALM	NERVOUS	1.862 NS	.527 NS
HAPPY	SAD	**6.697***	-.472 NS
ENTERTAINED	BORED	**4.925***	-1.979 NS
AT EASE	SCARED	1.996 NS	.671 NS
INQUISITIVE	DISINTERESTED	**10.615***	-4.031**
CONFIDENT	WARY	**4.717***	2.414*
FRIENDLY	UNFRIENDLY	**13.038***	**-6.994***
BRAVE	SCARED	**4.778***	1.739 NS
LAID-BACK	WORRIED	.488 NS	.326 NS
CALM	BOISTEROUS	-3.730**	-1.360 NS
AFFECTIONATE	COLD	**11.546***	**-8.384***
STEADY	JUMPY	-.703 NS	-1.005 NS
NOSY	NOT BOTHERED	**7.910***	-3.127*
STRONG	ANXIOUS	2.478*	2.391*
DOMINANT	SUBMISSIVE	**5.657***	**4.652***
ENERGETIC	LAZY	**7.395***	2.308*
RELAXED	UPTIGHT	1.348 NS	-1.036 NS
ALERT	UNALERT	3.808**	1.194 NS
GENTLE	HARD	1.426 NS	**-6.818***
COMFORTABLE	UNCOMFORTABLE	2.575**	-1.279 NS
QUIET	LOUD	-2.668*	-1.821 NS
EASY GOING	STUBBORN	.760 NS	**-3.947**
SLOW	FAST	**-4.948***	-3.103*

8.2.2.2.3. Correlation of dimensions

We correlated the personality scores produced by both familiar and unfamiliar observers for Stable 2. There was no significant correlation between personality scores generated by the two observer groups (Table 8.17).

Table 8.17

Spearman correlation coefficient for Stable 2

Spearman rho	**Dimension 1 Unfamiliar**	**Dimension 2 Unfamiliar**
Dimension 1 Familiar	.11 NS	.082 NS
Dimension 2 Familiar	.407 NS	.126 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

8.2.2.3. Stable 3

8.2.2.3.1. MDS

There were only familiar observers assessing the horses from Stable 3. The perceptual map (Figure 8.19) shows 12 horses positioned along the two dimensions. Gandalf is highest on Dimension 2, while Sophie is the lowest. Danny and Fleur are at the opposite ends on Dimension 1. As with the previous two stables the RSQ value of 0.91479 indicates that the high proportion of data is accounted for by the corresponding distances. The low stress value of 0.12851 shows good goodness of fit.

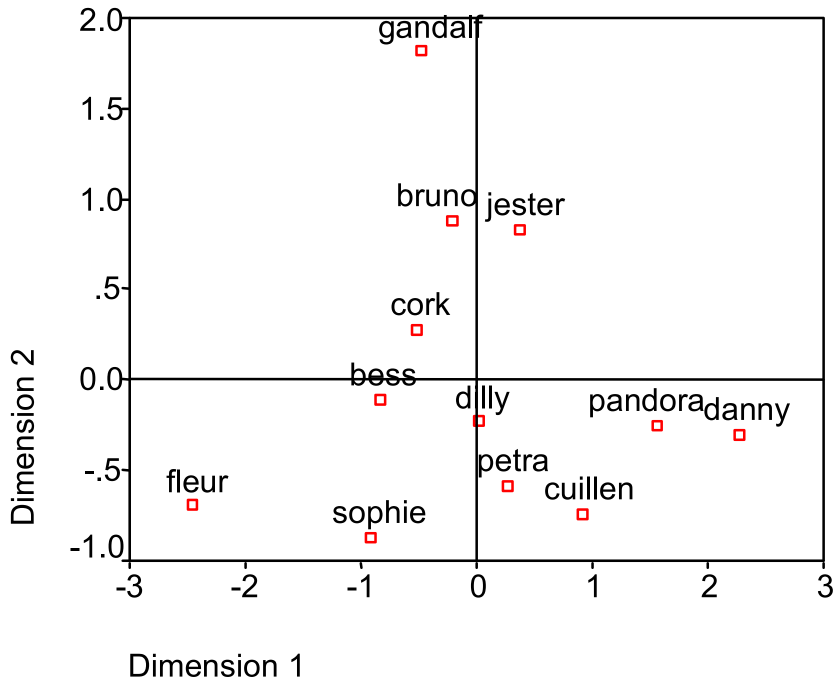


Figure 8.19

Euclidean two-dimensional distance model produced by familiar observers

8.2.2.3.2. Identification of dimensions

As with Stables 1 and 2, linear regression was used to associate the mean scores for each construct with two dimensions. Regression coefficients show that the following constructs are significantly associated with the first dimension: "confident-timid", "bold-shy", "nervous-bold", "relaxed-tense", "secure-insecure" and "brave-scared" (Table 8.18). These adjectives make up the *anxiety* facet of neuroticism (Costa & McCrae, 1992). Therefore we labelled the first dimension *Neuroticism* (N). The constructs significantly associated with the second dimension are "easy to work with-difficult", "gentle-rough", "patient-impatient" and "obedient-disobedient" (Table 8.17). They make up the *compliance* and *altruism* facets of the agreeableness dimension. The second dimension was therefore labelled *Agreeableness* (A).

Table 8.18

Linear regression to identify two dimensions - Stable 3 – Familiar observers

CONSTRUCTS		T – Regression coefficient	
		Stable 3	
		DIM 1	DIM2
FRIENDLY	UNFRIENDLY	-2.935*	-.842 NS
ATTENTION SEEKER	INDEPENDENT	**-3.754****	-.284 NS
CONFIDENT	TIMID	**-11.464*****	-.031 NS
BOLD	SHY	**-14.725*****	**3.750****
EASY TO HANDLE	DIFFICULT	-.898 NS	**-3.722****
SOCIABLE	UNSOCIABLE	-1.352 NS	-.017 NS
BOSSY	SUBMISSIVE	-2.539*	.651 NS
HAPPY	UNHAPPY	-4.101**	-1.377 NS
LAID BACK	HIGHLY STRUNG	-1.466 NS	-2.600*
CALM	EXCITABLE	-.896 NS	-2.231 NS
PLAYFUL	BORING	-3.158*	.904 NS
AFFECTIONATE	AGGRESSIVE	-.333 NS	-1.084 NS
OBEDIENT	DISOBEDIENT	1.078 NS	**-3.872****
NERVOUS	BOLD	**13.053*****	1.415 NS
INTERESTED	DISINTERESTED	-2.646*	.782 NS
AGGRESSIVE	GENTLE	-1.937 NS	2.164 NS
DOMINANT	SUBMISSIVE	-2.243 NS	2.023 NS
GENTLE	ROUGH	**5.356*****	**-5.136*****
EXPERIENCED	INEXPERIENCED	-2.784*	-2.219 NS
MATURE	IMMATURE	-1.835 NS	-3.035*
EASY TO WORK WITH	DIFFICULT	-1.074 NS	**-5.439*****
CHEEKY	WELL BEHAVED	-3.156*	1.001 NS
SECURE	INSECURE	**-9.807*****	**-4.552*****
NASTY	NICE	-2.403*	2.805*
BRAVE	SCARED	**-6.990*****	-.253 NS
FORWARD	HESITANT	**-5.339*****	.919 NS
BULLY	BULLIED	-1.575 NS	.966 NS
STUBBORN	WILLING	-.847 NS	2.264*
INTELLIGENT	THICK	-1.816 NS	1.514 NS
RELAXED	TENSE	**-8.162*****	**-4.574*****
CONTENT	GRUMPY	-2.387*	-.544 NS

PATIENT	IMPATIENT	.990 NS	**-4.620*** **
CLEVER	STUPID	-3.469**	1.230 NS
BARGY	QUIET	-4.348**	3.655**

P<0.05; ** P<0.01; *** P<0.001; NS - not significant.

8.2.2.4. Summary of identifying personality dimensions

We identified three personality dimensions, extroversion, neuroticism and agreeableness, across three different stables, three groups of familiar observers and one group of unfamiliar observers. Each group of familiar observers identified two personality dimensions. The first group of familiar observers clearly defined the agreeableness dimension. The same dimension was identified by the familiar observer in Stable 3. Extroversion was identified in Stables 1 and 2, but not in Stable 3. Neuroticism was identified only in Stables 2 and 3 (linear regression coefficients for all three familiar groups are included in Appendix 24). Unfamiliar observers generated two dimensions, extroversion and neuroticism. The dimensions overlapped in both stables. In the first stable some constructs associated with the neuroticism dimension were also associated with extroversion. In the second stable the first group of constructs clearly constituted the extroversion dimension. The second group of constructs, however, could be associated with either extroversion or neuroticism; therefore, the second dimension was labelled extroversion/neuroticism. The linear regression coefficients for unfamiliar observers, for both stables, are included in Appendix 24.

8.2.3. Horses' performance scores

Familiar observers, who knew the horses well and had worked with them over a longer period of time, were asked to rate the horses on their overall performance in most situations (for example, how efficient the horse is for riding, how cooperative it is etc.). Each horse was rated on a nine point scoring scale. The observers had to circle 1 if they thought the horse's overall working performance was very poor and circle 9 if they thought the horse's overall performance was excellent (an example of the performance scale is included in Appendix 19). The results show significant agreement between the observers in each observer group.

In this section we also look at the personality profiles of individual horses that scored high or low on performance scale.

8.2.3.1. Stable 1

Familiar observers from Stable 1 agreed significantly on rating their horses on a performance scale (W=0.242; p=0.026). Barney and Blue got the highest mean performance score; Zak and Gunther got the lowest scores at 5.88 (Figure 8.20).

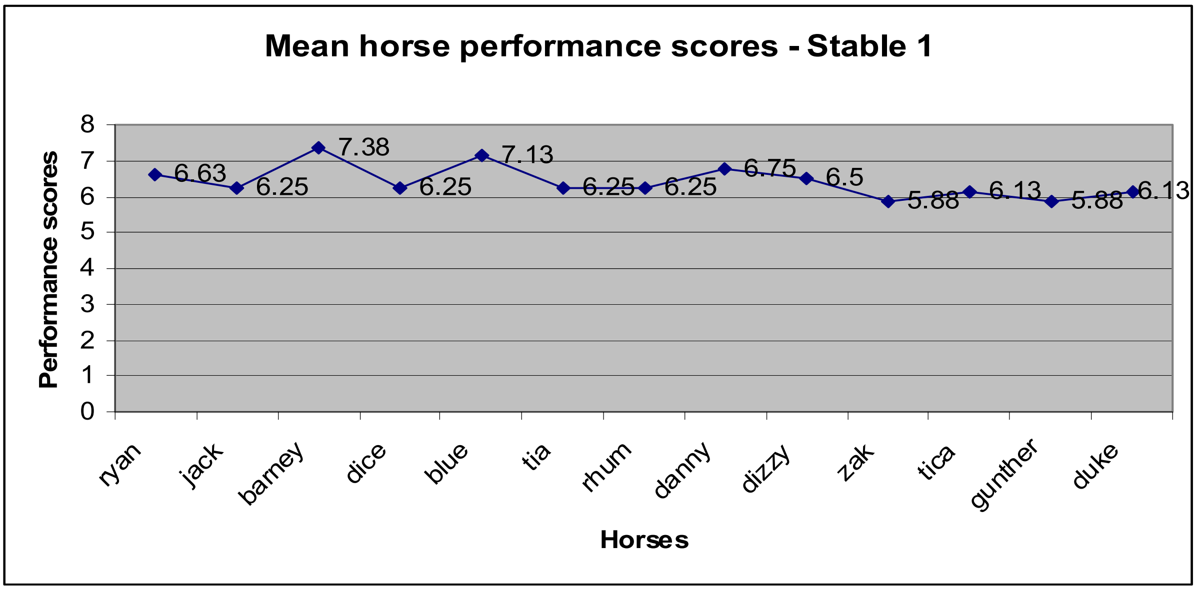


Figure 8.20

Mean horse performance scores Stable 1

They were perceived as the two most agreeable horses (see Figure 8.15). Looking at the personality profiles, both horses tend to have low ratings for most of the constructs. The scoring list (see Table 8.5, page 149-150, for the list of constructs) shows that lower ratings are mostly associated with the construct poles indicative of agreeableness. For example Barney scored very low on constructs five ("easy to handle – difficult") and 21 ("easy to work with – difficult"). At the same time he scored high on constructs 24 ("nasty – nice") and 34 ("bargy – quiet") (Figure 8.21). That means that he was perceived as a nice horse who is easy to handle and work with.

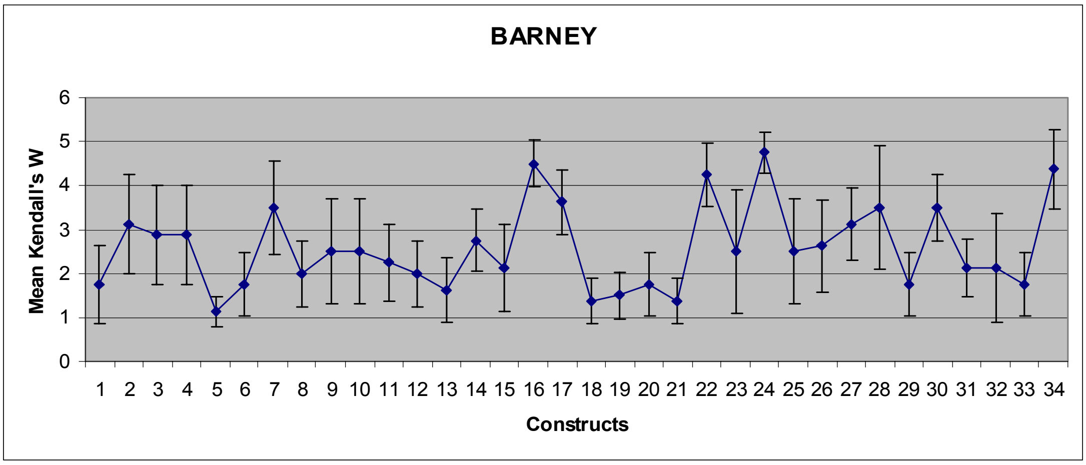


Figure 8.21

Personality profile for Barney – high on performance scale

Blue, who was also a high performer, scored low on constructs four ("bold – shy") and 15 ("interested – disinterested"). He also got high ratings for constructs 16 ("aggressive – gentle"), 24 ("nasty – nice"), 28 ("stubborn-willing"), and 32 ("patient-impatient") (Figure 8.22). Therefore Blue is an interested and bold horse that is also impatient, nice and willing.

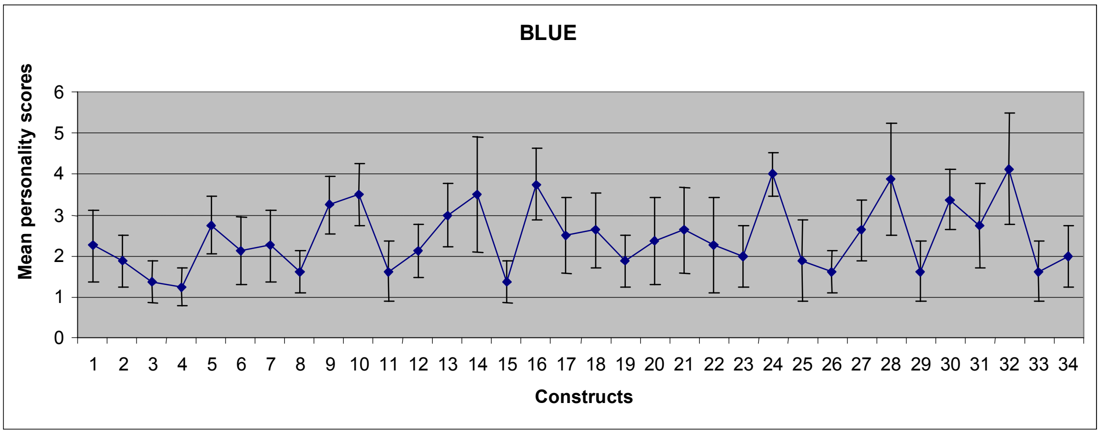


Figure 8.22

Personality profile for Blue – high on performance scale

Zak and Gunther were horses that had the lowest performance scores. In contrast with Barney and Blue, Zak scored high on constructs five ("easy to handle – difficult") and 21 ("easy to work with – difficult"). So he was perceived as a difficult horse to handle and to work with. He also had low scores for constructs 26 ("forward – hesitant") and 34 ("bargy – quiet"), so he was also perceived as forward and bargy or pushy (Figure 8.23).

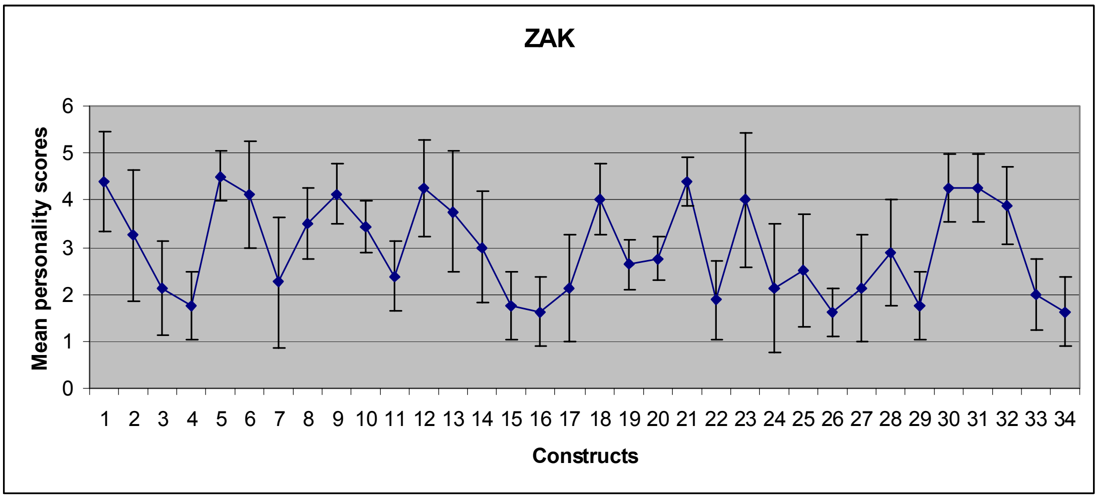


Figure 8.23

Personality profile for Zak – low on performance scale

Gunther, like Barney, had high scores for constructs 24 ("nasty – nice") and 34 ("bargy–quiet"). So he was seen as nice and quiet. At the same time he scored low on constructs nine ("laid back – highly strung"), 10 ("calm-excitable") and 21 ("easy to work with – difficult") (Figure 8.24). So he was also seen as laid back, calm and easy to work with. However, his performance was still low.

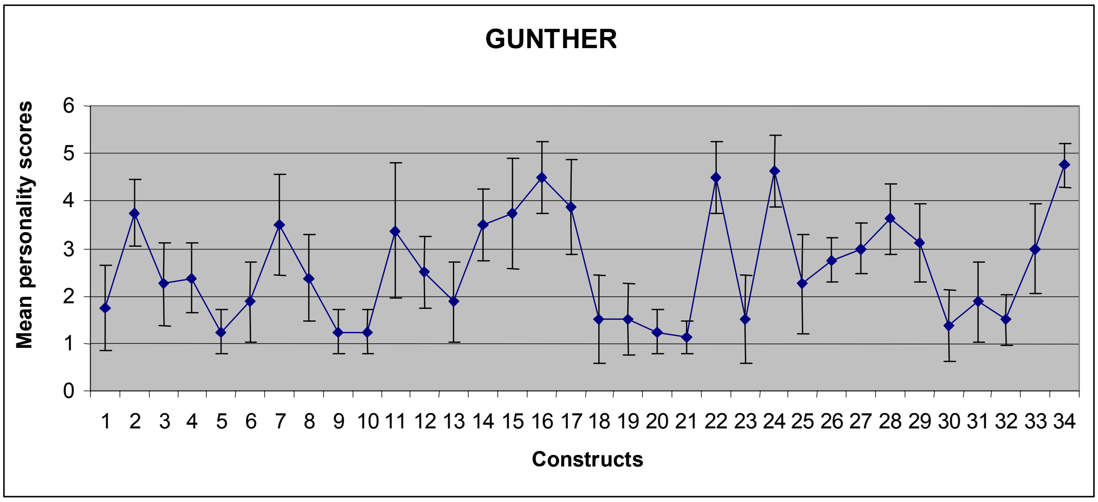


Figure 8.24

Personality profile for Gunther – low on performance scale

8.2.3.2. Stable 2

The observers from the second stable also agreed significantly on how well their horses performed ($W=0.523$; $p<0.001$). Rocky and Jazz were rated highest on the performance scale, Lil and Kurt lowest (Figure 8.25). The perceptual map (see Figure 8.15) shows that Kurt and Lil were positioned close to each other and were the lowest on neuroticism (Dimension 1), which means they were both nervous and insecure. Rocky and Jazz were also located very near each other but at the opposite end of the perceptual map to Lil and Kurt. Rocky and Jazz were high on Dimension 1 (neuroticism), which means they were emotionally more stable. They were also lower on Dimension 2 (extroversion), meaning they were more friendly, gentle and sociable.

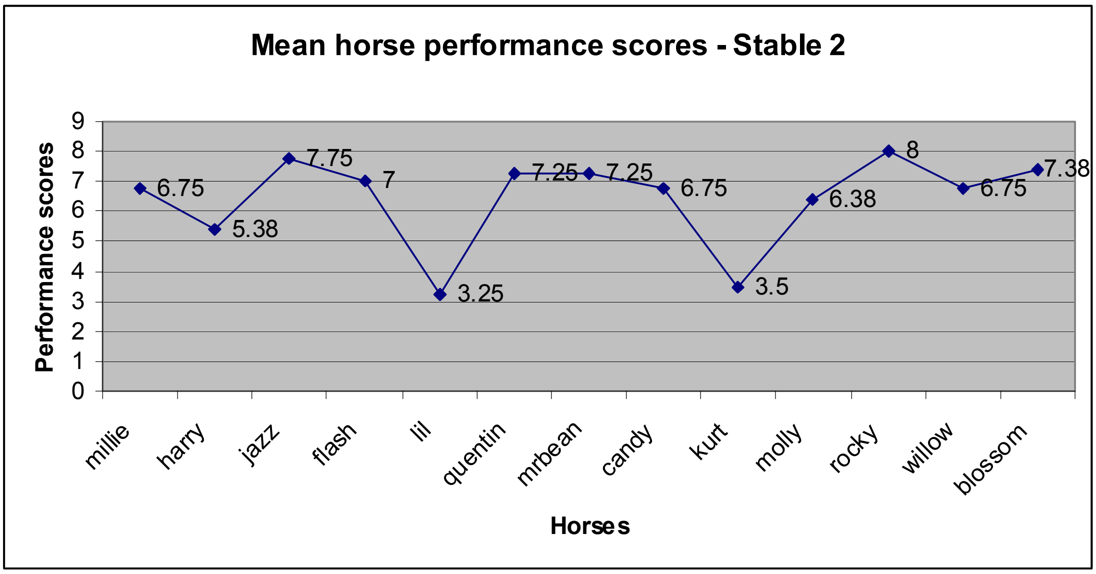


Figure 8.25

Mean horse performance scores Stable 2

Rocky was rated low on constructs five ("easy to handle-difficult"), 21 ("easy to work with-difficult") and 25 ("brave-scared"). He also received high scores for constructs 16 ("aggressive-gentle") and 24 ("nasty-nice"). Therefore he was perceived as a nice and gentle horse that is easy to handle and to work with.

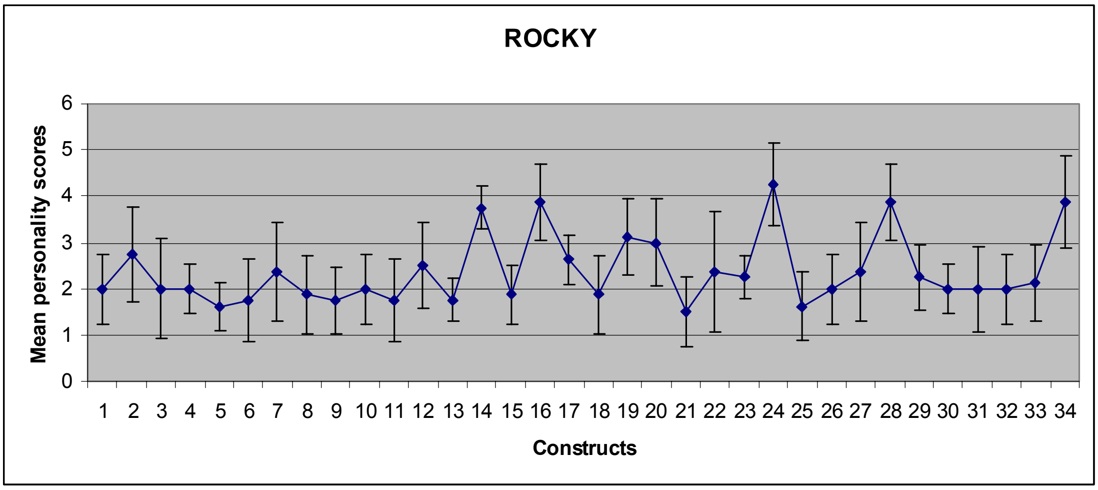


Figure 8.26

Personality profile for Rocky – high on performance scale

There is more variation in Jazz's personality profile. He has extremely low scores on constructs five ("easy to handle-difficult"), 10 ("calm-excitable"), 18 ("gentle-rough" and 21 ("easy to work with-difficult"). He also has extremely high scores on constructs 16 ("aggressive-gentle") and 24 ("nasty-nice") (Figure 8.27). The familiar observers therefore agreed on Jazz being a nice, gentle and calm horse that is easy to handle and to work with.

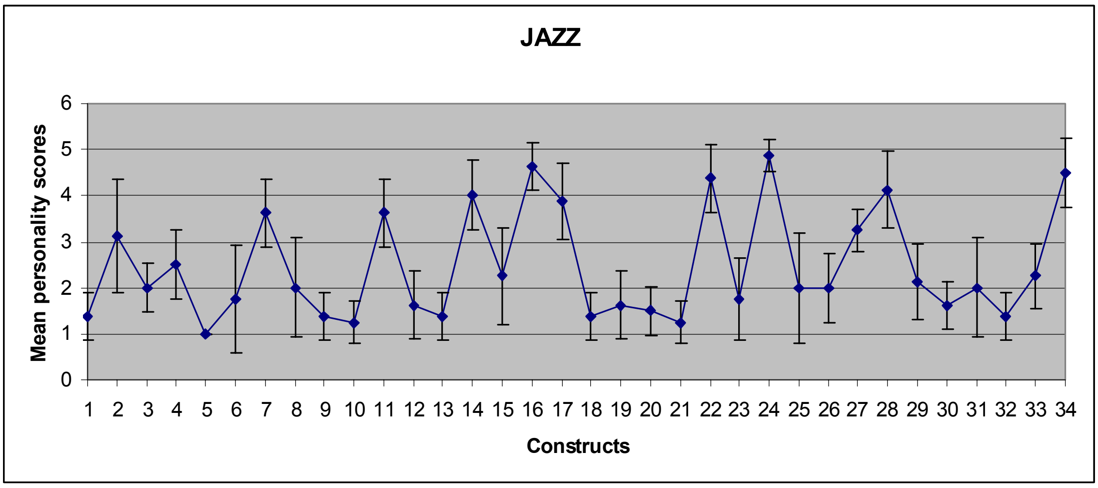


Figure 8.27

Personality profile for Jazz – high on performance scale

The lowest performance scores were received by Lil and Kurt. Lil received high scores on constructs three ("confident-timid"), 16 ("aggressive-gentle"), 19 ("experienced-inexperienced"), 25 ("brave-scared"), 26 ("forward-hesitant") and 30 ("relaxed-tense"). She received low scores on constructs one ("friendly-unfriendly"), 12 ("affectionate-aggressive"), 14 ("nervous-bold") and 18 ("gentle-rough") (Figure 8.28). Therefore she was perceived as timid, friendly, affectionate and gentle, but at the same time as a horse who is very inexperienced, tense, nervous, scared and hesitant.

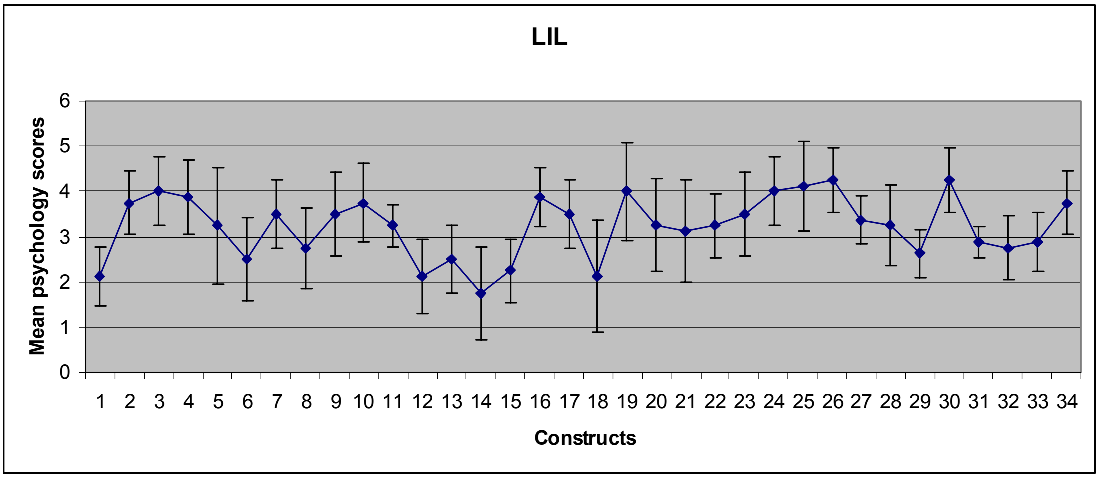


Figure 8.28

Personality profile for Lil – low on performance scale

Kurt was also a low performer. However, his personality profile is a bit different from Lil's. The observers rated him high on constructs six ("sociable-unsociable"), 11 ("playful-boring"), 19 ("experienced-inexperienced") and 20 ("mature-immature"); and low on constructs 13 ("obedient-disobedient") and 18 ("gentle-rough") (Figure 8.29). Even though he was perceived as obedient and gentle, he was also rated as a horse who is unsociable, boring, inexperienced and immature.

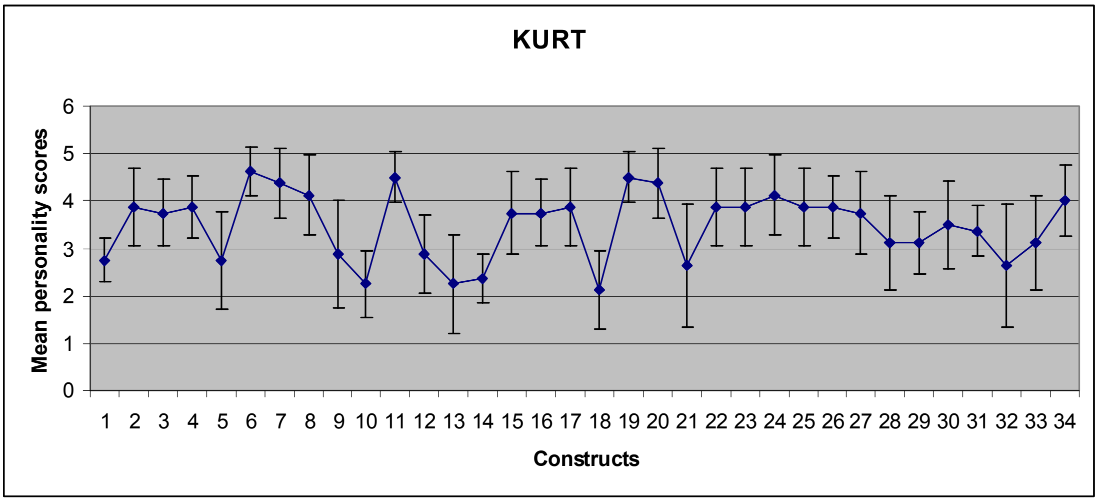


Figure 8.29

Personality profile for Kurt – low on performance scale

8.2.3.3. Stable 3

Observers from Stable 3 rated Bruno and Jester as the best performers. Sophie and Fleur were perceived as the poorest performers (Figure 8.30). Kendall's W shows that the observers in Stable 3 also agreed significantly on how they rated their horses (W=0.375, p<0.001). Both Bruno and Jester were quite high on the agreeableness dimension and positioned nearby on MDS perceptual map. Similarly, Sophie and Fleur were located moderately close, however, they were both low on the agreeableness dimension and far away from Bruno and Jester (see Figure, 8.17).

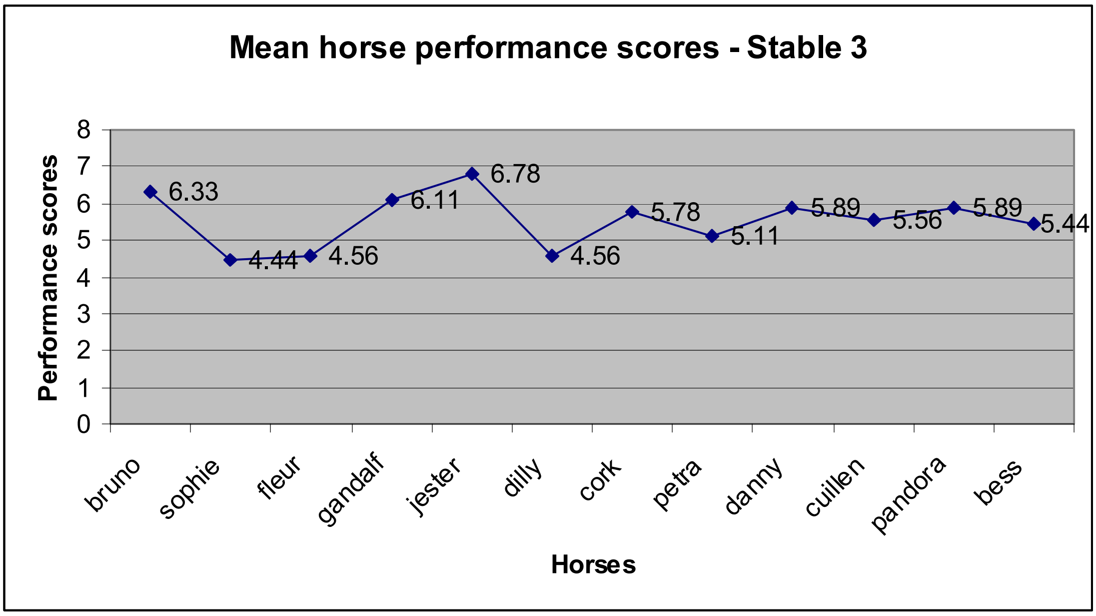


Figure 8.30
Mean horse performance scores Stable 3

Bruno, a high performer, scored high on constructs 16 ("aggressive-gentle") and 24 ("nasty-nice"), and low on constructs 15 ("interested-disinterested"), 18 ("gentle-rough"), 22 ("cheeky-well behaved") and 32 ("patient-impatient") (Figure 8.31). So, Bruno was perceived as a nice, gentle and patient horse, who was also cheeky and interested.

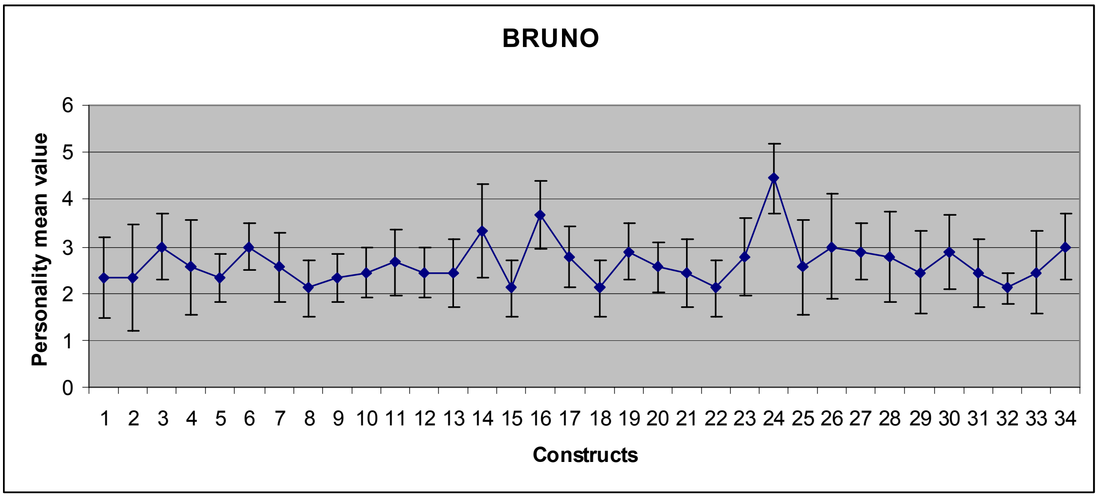


Figure 8.31

Personality profile for Bruno – high on performance scale

Jester, also a high performer, was high on constructs 2 ("attention seeker-independent"), 14 ("nervous-bold"), 24 ("nasty-nice") and 34 ("bargy-quiet"), and low on constructs 7 ("bossy-submissive") and 27 ("bully-bullied") (Figure 8.32). Therefore Jester is independent, bold, bossy and a bully, but also a horse who is nice and quiet.

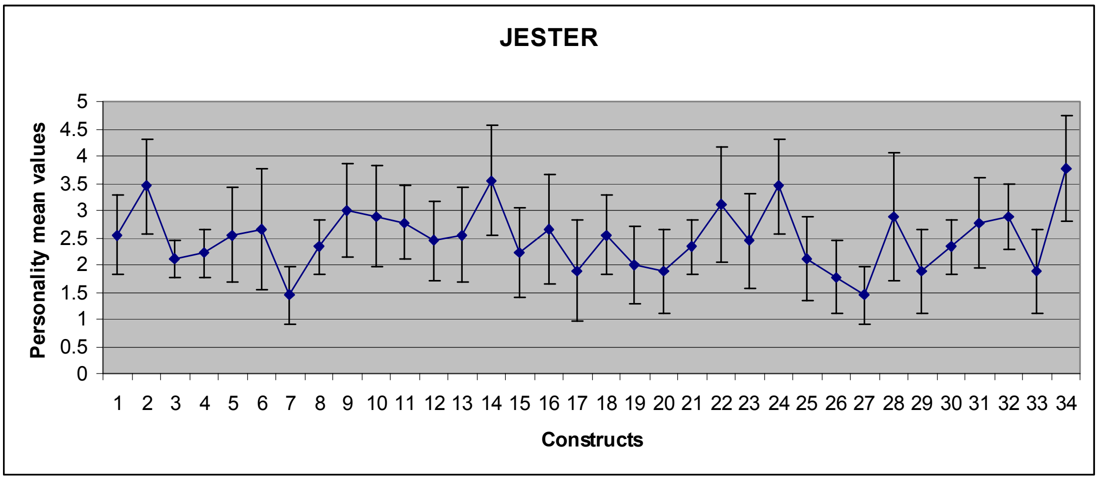


Figure 8.32

Personality profile for Jester – high on performance scale

The familiar observers rated Sophie and Fleur as the poorest performers. Sophie scored high on constructs 9 ("laid back-highly strung"), 10 ("calm-excitable"), 21 ("easy to work with-difficult") and 32 ("patient-impatient"). She also scored low on constructs 29 ("intelligent-thick") and 33 ("clever-stupid") (Figure 8.33). Therefore, although she was rated as intelligent and clever, she was also seen as highly strung, excitable, impatient and difficult to work with.

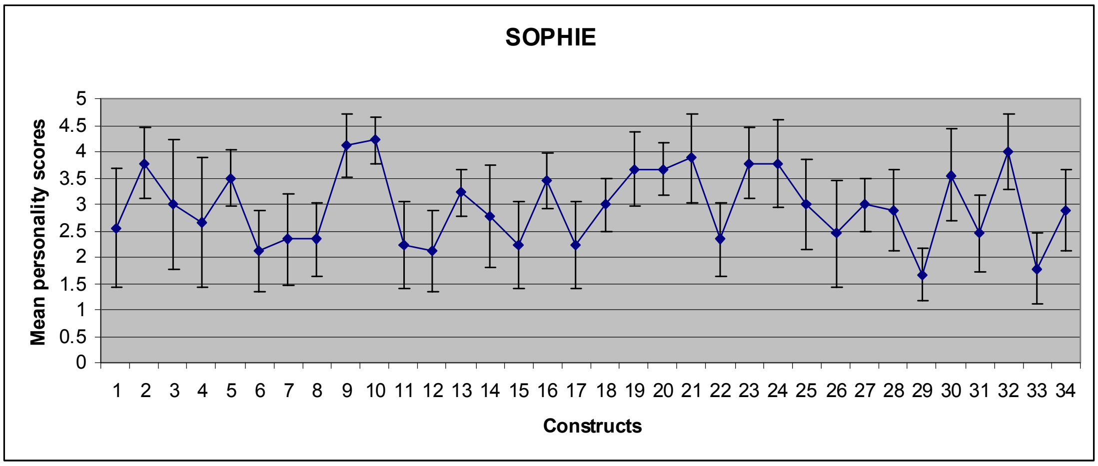


Figure 8.33

Personality profile for Sophie – low on performance scale

The observers rated Fleur as a low performer as well. She was high on constructs 19 ("experienced-inexperienced") and 20 ("mature-immature"), and low on constructs 14 ("nervous-bold"), 15 ("interested-disinterested"), and 29 ("intelligent-thick"). Like Sophie, Fleur was also rated as intelligent and interested, but unlike Sophie, Fleur was scored as an immature and inexperienced horse.

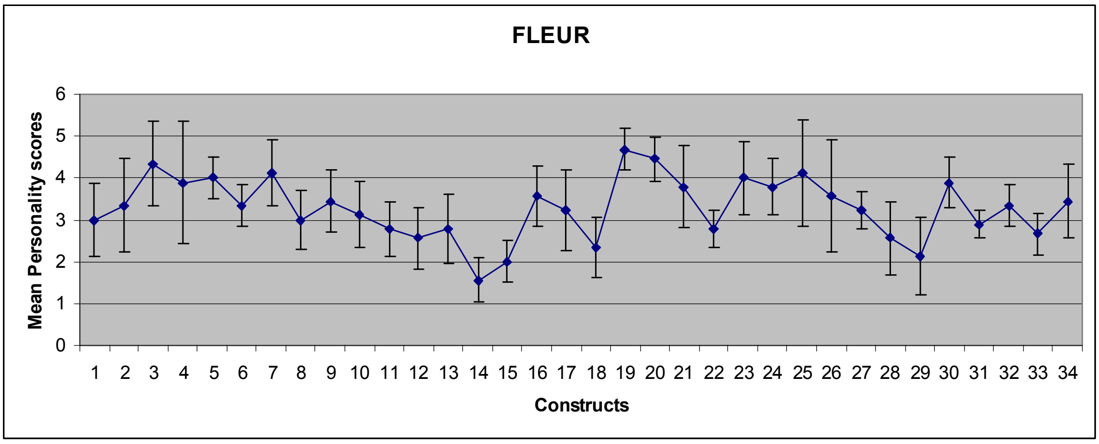


Figure 8.34

Personality profile for Fleur – low on performance scale

To sum up, horses with high and low performance vary in their personality profiles. All high performers were rated as nice. Barney, Blue, Rocky and Jazz were also rated as easy to handle and to work with. Rocky, Jazz and Bruno were perceived as gentle. Bruno was also seen as a patient horse, but at the same time cheeky and interested. Blue and Jester were both independent and bold. Blue was also impatient and willing, and Jester was bossy and a bully.

As an example of a horse with low performance scores, Gunther was quite atypical as he was scored as nice and quiet, laid back, calm and easy to work with. However, most low performers were rated as difficult to handle and to work with. Lil, Kurt and Fleur were also rated as inexperienced and immature. Even though Lil was gentle and friendly, she was also seen as tense, nervous, scared and hesitant. Kurt, even though obedient and gentle, was rated as unsociable and boring. Fleur was inexperienced, but rated as intelligent and interested. Sophie was likewise perceived as intelligent and clever, but, similarly to Lil, highly strung, excitable and impatient.

Looking at the best and the poorest performers in each stable, we can conclude that it is not only one construct or personality trait that contributes to a horse being rated as a good or bad performer. It is rather a combination of personality traits that are responsible for horses' performance.

Having explored the performance and personality profiles of individual horses, we will now correlate performance scores and personality dimensions, discovered with MDS.

8.2.3.4. Horse performance and personality dimensions

As there were some constructs consistently associated with high or low performance, we expected a relationship between horses' performance and the two personality dimensions generated with MDS.

Correlation of performance and personality scores showed there was no relationship between performance and agreeableness in Stable 1 ($p>0.05$). There was also no correlation between performance and extroversion ($p>0.05$) (Table 8.19).

Table 8.19

Correlation of horse performance and personality dimensions – Stable 1

		Dimension 1 – Agreeableness	Dimension 2 – Extroversion
Performance	**Spearman's rho**	**0.182 NS**	**0.539 NS**
	P	**0.553**	**0.057**

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

In Stable 2 there was a significant correlation between performance scores and the neuroticism dimension ($p<0.05$). However, the correlation between performance and extroversion was not significant ($p>0.05$) (Table 8.20).

Table 8.20

Correlation of horse performance and personality dimensions – Stable 2

		Dimension 1 – Neuroticism	Dimension 2 – Extroversion
Performance	**Spearman's rho**	**0.664***	**0.542 NS**
	P	**0.013**	**0.056**

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

In Stable 3 there was a highly significant correlation between performance scores and the agreeableness dimension (p<0.01). There was, however, no association between performance and neuroticism (p>0.05) (Table 8.21). In Stable 3 the agreeableness dimension was more clearly defined than in Stable 1, where it was mixed with adjectives which form facets of neuroticism.

Table 8.21

Correlation of horse performance and personality dimensions – Stable 3

		Dimension 1 – Neuroticism	**Dimension 2 – Agreeableness**
Performance	**Spearman's rho**	**0.488 NS**	**0.733****
	P	**0.108**	**0.007**

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

On the whole, there was a significant correlation between performance and neuroticism in Stable 2 and performance and agreeableness in Stable 3. There was no other significant correlation between performance and personality dimensions. We also correlated Kendall's coefficient of concordance W and performance scores to explore whether the observers agree more on horses that are rated as good performers. The results indicate that there was a significant correlation between the two scales for horses from Stable 2 (p<0.05), but not for horses from Stables 1 and 3 (p>0.05). The full report of the results is included in Appendix 25.

8.3. Synopsis

In this Chapter we have presented the results from the second study. The personality and performance of 38 horses, aged three to 21 years were assessed by 24 familiar and 10 unfamiliar observers. The age of participants, all female, varied from 17 to 32 years. We used Repertory grid technique as the method for data collection. Some changes were made in comparison to the first study: elicitations and interviews were done individually, while the scoring was done by all the observers in a group at the same time. The video part included a standardized novel object test, an open umbrella test. The purpose of this additional condition was to see if the correlation of horse personality scores between familiar and unfamiliar observers would improve.

We first calculated the general degree of agreement for each stable for familiar and unfamiliar observers. The degree of agreement was mostly significant but the familiar had significantly higher agreement than the unfamiliar observers. Elicitation was an important part of the method and the observers were expected to reach a higher degree of agreement on constructs that they had generated themselves. The Kendall's W was calculated separately for constructs elicited by at least half of the participants in the observer group and the rest. The results showed no consistent tendency towards a higher degree of agreement in any of the two groups of constructs. The first Stable had a higher degree of agreement when constructs were provided. Observers in Stable 3 showed a slightly higher degree of agreement on elicited constructs. We intended to further investigate whether the provision of constructs has an effect on the agreement. Kendall's W was calculated for the first and the last three constructs in each observer groups. The results did not show any consistency. In Stable 1 the mean Kendall's W was higher for the last three constructs in both observer groups. Familiar observers in Stable 2 had a slightly higher degree of agreement when rating horses on the first three constructs. However that was not the case for either unfamiliar observers or familiar observers in Stable 3. To explore whether the observers used their constructs consistently, we calculated Kendall's W for each construct separately. Significant Kendall's W indicated that the observers applied the constructs systematically. Familiar and unfamiliar observers each had their own scoring list which they had generated.

One of the central questions of the thesis and this chapter was how many and which personality dimensions horses have. Multidimensional scaling discovered two dimensions produced by both observer groups, familiar and unfamiliar. They were labelled neuroticism and extroversion. The familiar observers also generated constructs that were grouped in the agreeableness dimension.

In comparison to the first study, we added a novel object test to the video observation to investigate whether this would contribute to a higher correlation between the horse personality scores produced by familiar and unfamiliar observers. This happened in Stable 1, but not in Stable 2.

Finally, we calculated the performance scores for individual horses and compared them to the individual personality profiles. It seems that not only one construct, but certain combinations of constructs or personality traits are associated with different levels of performance. Generally, horses that are easy to lead and to work with also scored high on the performance scale. To explore any consistent patterns between personality and performance we correlated horse personality dimensions with performance scores. There was a significant correlation between performance and neuroticism in Stable 2 and performance and agreeableness in Stable 3.

CHAPTER 9

Observers' Personality, Degree of Empathy and Emotional Intelligence

9.1. STUDY 1

In this study we will first present descriptive statistics about the observers, their personality and empathy. We will then investigate whether observers' empathy and personality scores influence how they perceived and assessed the personality of the horses. The analysis was limited by the low number of observers in each observer group. The observers were treated as four different groups, each with their own scoring list (see Figure 7.1).

9.1.1. Methods

9.1.1.1. Participants

Four groups of observers rated two groups of horses from two stables. There were 44 participants and 21 horses (see section 7.1.1. for further details). The two familiar groups consisted of stables staff who worked with the horses. In the third group were people in general familiar with horses but not with the horses included in the study, psychology and vet students, labelled as experienced, and in the fourth group there were inexperienced raters without any experience with horses. Familiar observers assessed only their own horses, i.e. staff from Stable 1 assessed only their own horses and staff from Stable 2 assessed only their own horses. Experienced and inexperienced observers assessed horses from both riding schools (see Figure 7.1). The familiar observers rated the horses on the basis of their previous experience with the horses, while the experienced and inexperienced based their ratings on the video clips of the individual horses. We expected that observers' individual differences would influence their rating of horse personality.

9.1.2. Results

The observers completed three questionnaires about themselves: The Observer information sheet (the questionnaire is included in Appendix 26), NEO personality test (Costa & McCrae, 1992) and Animal Empathy Scale (AES) (Paul, 2000) (the AES is included in Appendix 27). Table 9.1 shows the mean values with standard deviations for observers' personality scores, Animal empathy scores, their age and the time they have worked with horses.

Table 9.1
Descriptive statistics for all the participants in Study 1

	Age (mean)	Animal Empathy Scale (mean)	Agree-ableness (mean)	Conscien-tiousness (mean)	Extrover-sion (mean)	Neuroti-cism (mean)	Openness (mean)	Experi-ence with horses (in years)
Stable 1 Familiar	44.75	138.00	27.25	36.50	25.25	20.25	23.50	15
Stable 2 Familiar	23.00	137.00	27.25	33.50	28.75	17.75	26.75	2.5
Experi-enced observers	20.79	147.05	32.47	30.47	33.00	25.58	32.05	10
Inexperi-enced observers	41.76	142.24	36.35	34.06	28.18	18.29	35.12	
Overall mean	31.27	143.45	33.02	32.68	30.05	21.57	31.98	
Standard deviation	12.17	22.04	5.72	7.64	6.89	9.58	7.22	

9.1.2.1. Observers' information sheet

The rater information form was used to provide background information about the participants, which could be relevant in determining the degree of agreement and horse personality dimensions. Age, level of education, whether they were vegetarians or not, and how long they had been familiar with horses was the information requested. Level of education and vegetarianism were not included as variables in the analysis due to the low number of participants (This information is included in Appendix 28).

The mean age of the observers was 31.5 years. There was variation in age between the three observer groups (Figure 9.1). The inexperienced group was the oldest (mean=41.8, SD=7.5). The youngest were the experienced observers (mean=20.8, SD=0.9). The greatest variation was in the familiar groups (SD=13.5) where the mean age was 33.9 years. The experienced observers were significantly younger than both the familiar ($t=2.917$, $p<0.05$) and inexperienced observers ($t= 10.411$, $p<0.001$).

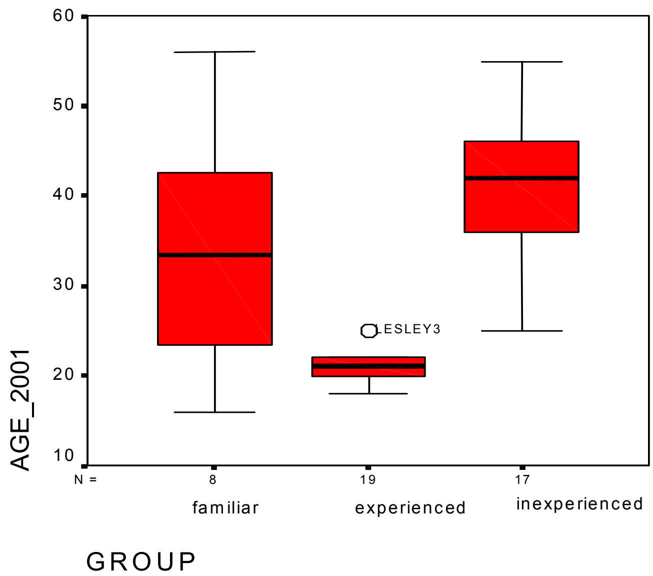


Figure 9.1

Boxplot showing the age of observers

9.1.2.2. Animal Empathy Scale (AES)

The Animal Empathy Scale (Paul 2000) was developed based on the Mehrabian and Epstein Emotional empathy scale (1972). It indicates observers' degree of empathy towards animals in general. The three observer groups had a similar distribution of empathy scores (Figure 9.2).

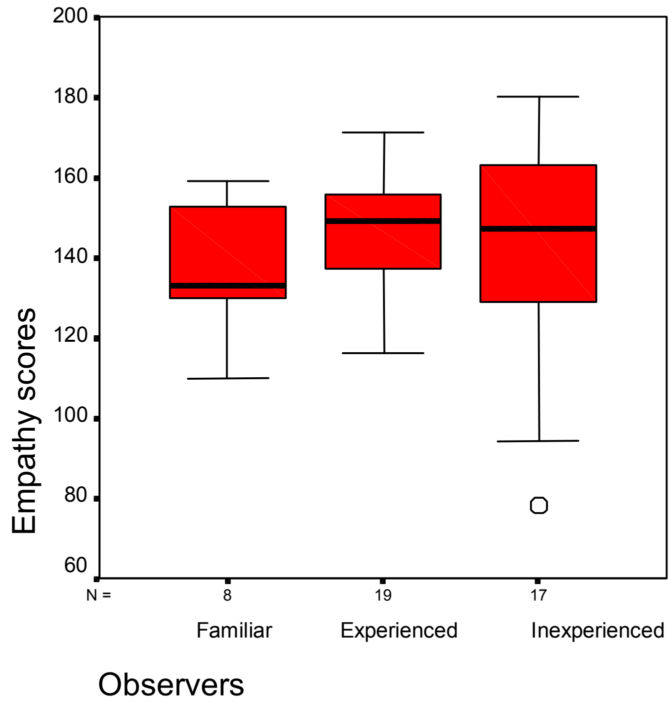


Figure 9.2
Animal Empathy Scale scores for the three observer groups

ANOVA showed there was no significant difference in the scores for empathy between the three groups ($F=0.560$; $P>0.05$).

9.1.2.3. NEO Five factor personality questionnaire

Observers completed the NEO five factor personality questionnaire (Costa & McCrae, 1992). Figure 9.3 shows the distribution of personality scores for the three observer groups.

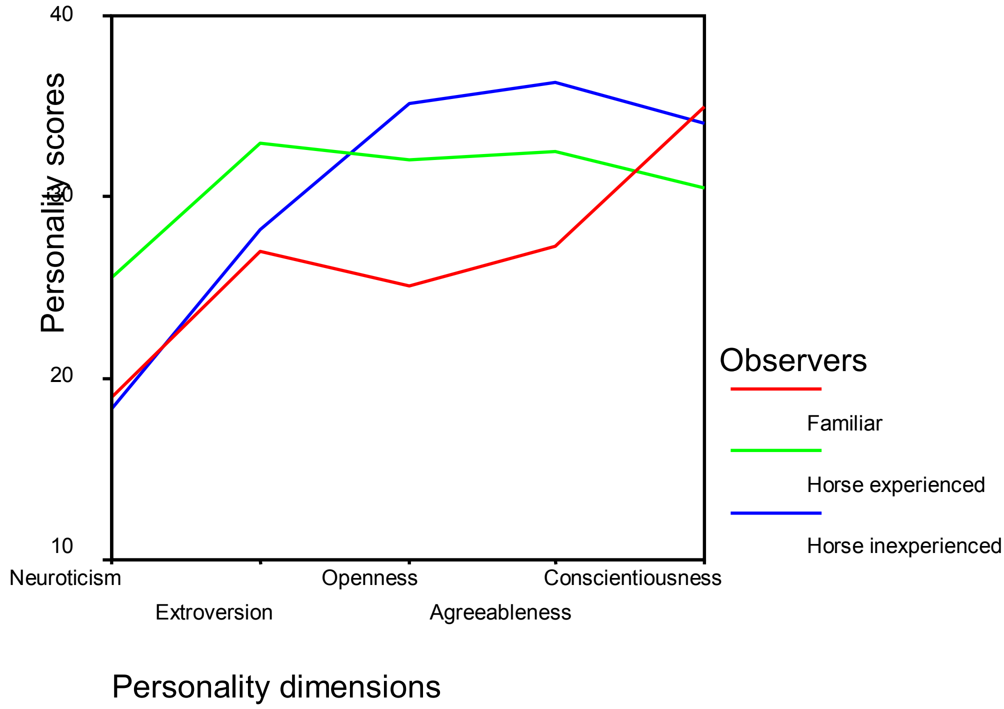


Figure 9.3

Distribution of personality scores for three observer groups.

ANOVA was performed on each of the five dimensions for the three observer groups. It showed a significant difference in the scores for the neuroticism dimension between the three groups ($F=3.258$; $P<0.05$). A further T-test showed that the experienced observers had significantly higher N scores than the inexperienced ($t=2.268$; $p<0.05$). ANOVA also showed there was significant difference in the scores for the extroversion dimension between the three groups ($F=3.527$; $P<0.05$). Experienced observers scored significantly higher than familiar ($t=-2.575$; $P=0.016$) and inexperienced observers ($t=2.171$; $P=0.037$). There was a significant difference in the openness scores between the three groups (F=6.554; P=0.003). Familiar observers scored significantly lower than inexperienced (t=-3.176; P=0.004) and experienced observers (t=2.171; P=0.037).

There was also a significant difference in the scores for the agreeableness dimension between the three groups (F=10.007; P=0.000). Inexperienced observers scored significantly higher than experienced (t=-2.425; P=0.021) and familiar (t=-4.641; P=0.000). Finally, there was no significant difference in the scores for the conscientiousness dimension between the three groups (F=1.470; P=0.242).

9.1.2.4. Difference in the degree of agreement between participants who scored low and high on the Animal Empathy Scale

Does observers' ability to empathise influence their degree of agreement when assessing horse personality?

The median scores for females on the Animal Empathy Scale (AES) developed by Paul (2000) was (n=278) 142. In this study the median for all the participants (females only) was 147 and the mean 143. Participants in each observer group (familiar, experienced and inexperienced) were divided into two groups: those whose overall score was lower than 147 (median) and those who scored higher than 147 on the Animal Empathy Scale. We were interested to see if there was any difference in the degree of agreement between the observers who scored low and those who scored high on the empathy scale.

There were not enough familiar observers to split them into high and low empathy groups, although we had almost equal numbers of unfamiliar observers high and low on the Empathy scale. We therefore calculated the degree of agreement separately for observers high and low on the AES for both stables.

9.1.2.4.1. Stable 1

Kendall's W was calculated for both unfamiliar observer groups, experienced and inexperienced. There were 10 experienced participants who scored high and nine who scored low on the AES (Figure 9.4). In the inexperienced group there were eight who scored high and nine who scored low on the AES (Figure 9.5).

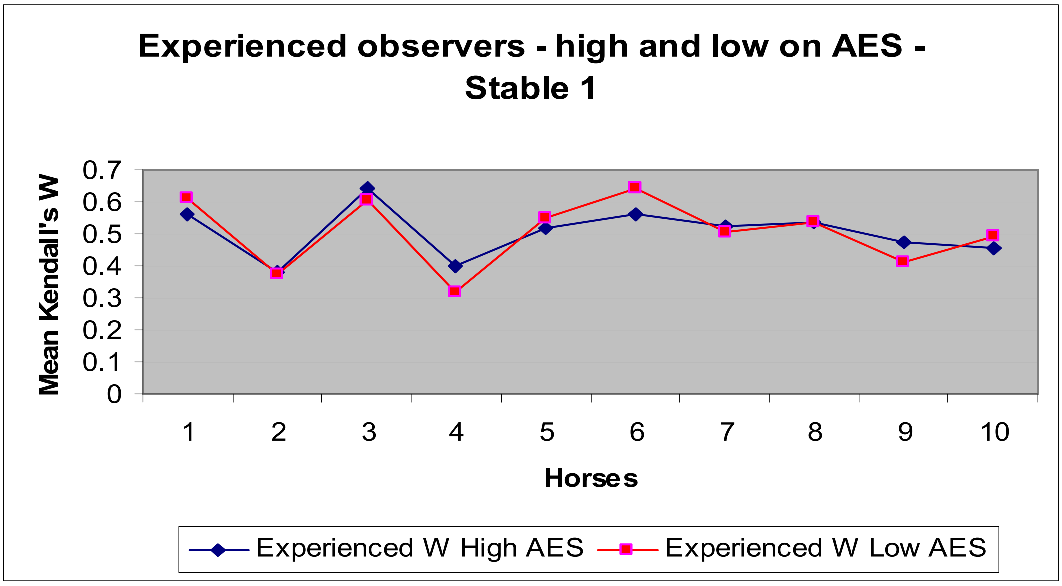


Figure 9.4

Experienced observers, high and low on the empathy scale, Stable 1

There seems to be more variation between high and low empathic inexperienced observers (a table with exact Kendall's W values is included in Appendix 29).

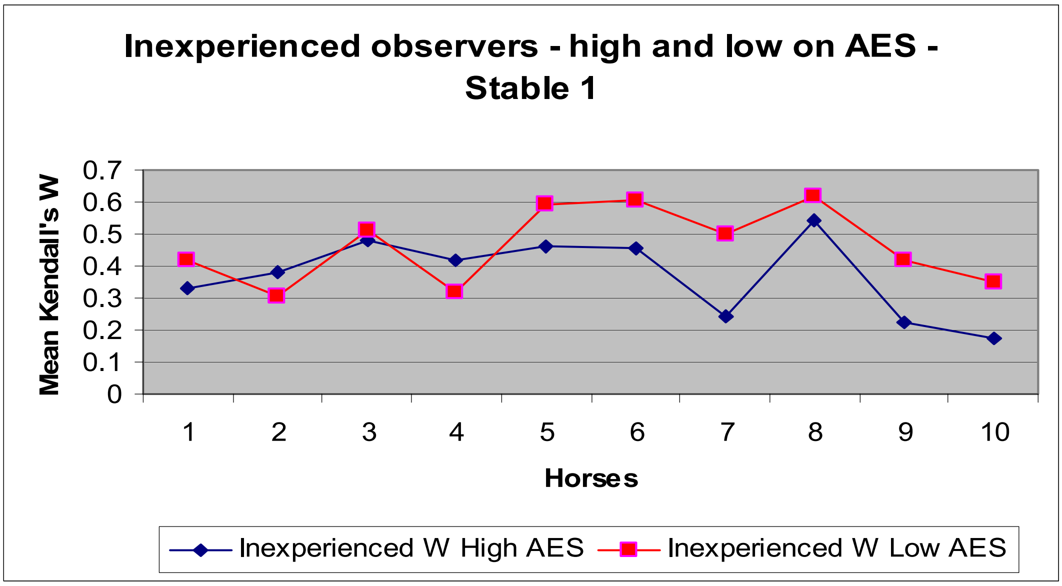


Figure 9.5

Inexperienced observers, high and low on the empathy scale, Stable 1

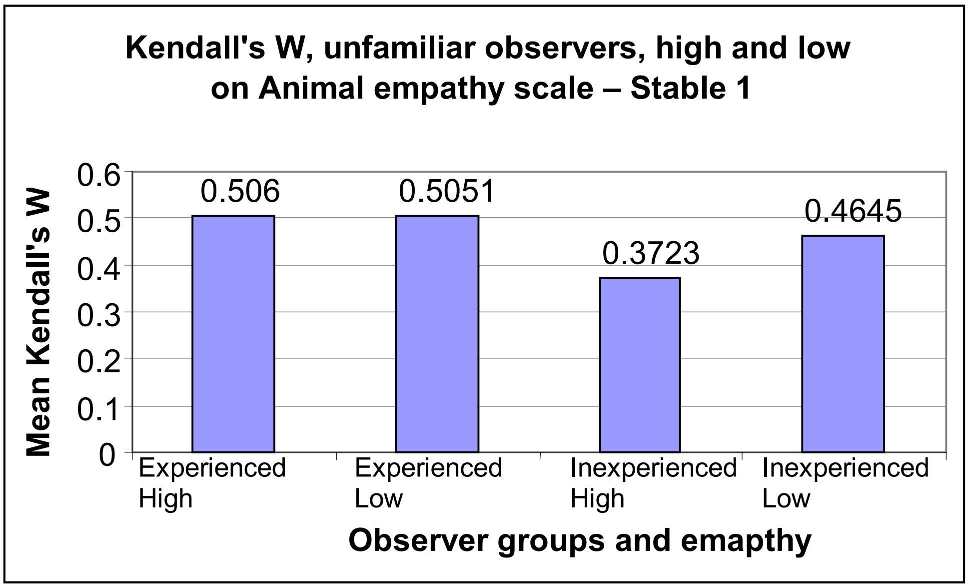


Figure 9.6

Kendall's W mean values per group, for high and low AES, Stable 1

There was not much difference in the mean values produced by experienced observers high and low on the AES (Figure 9.6). The t-test showed no significant differences between less and more empathic experienced observers; however, less empathic inexperienced observers agreed significantly more ($t=2.584$, $p<0.05$).

9.1.2.4.2. Stable 2

To see if there was any difference in the degree of agreement between participants who scored high and low on the empathy scale Kendall's W was performed for each group of observers separately (a table with the exact Kendall's W values for Stable 2 is included in Appendix 29).

There seem not to be any consistent trends in how low or high empathic observers agreed on horse personality scores, either in the experienced (Figure 9.7) or the inexperienced (Figure 9.8) group of observers.

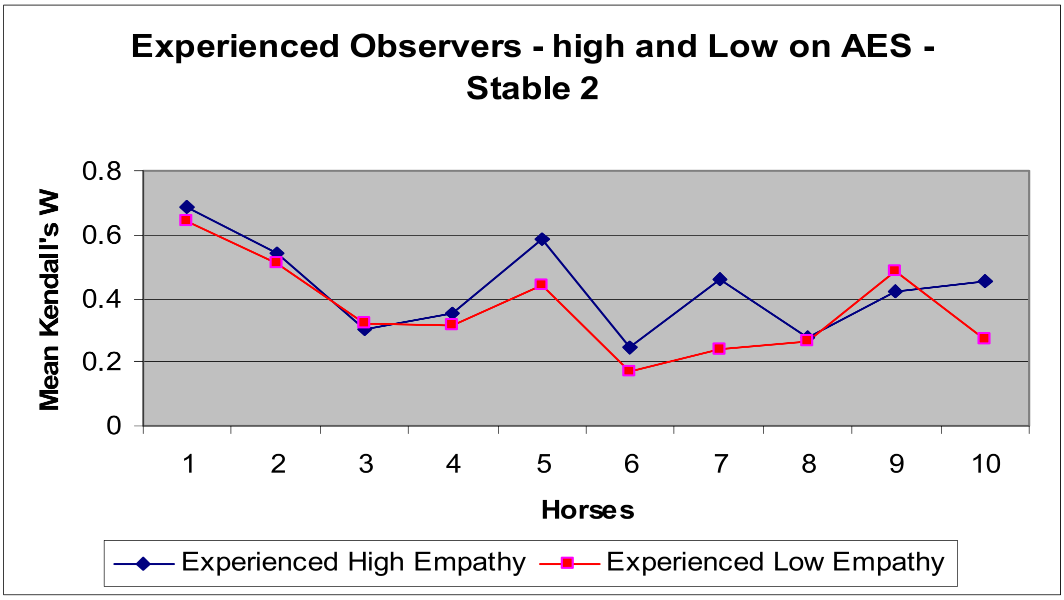


Figure 9.7

Kendall's W for experienced observers high and low on the empathy scale

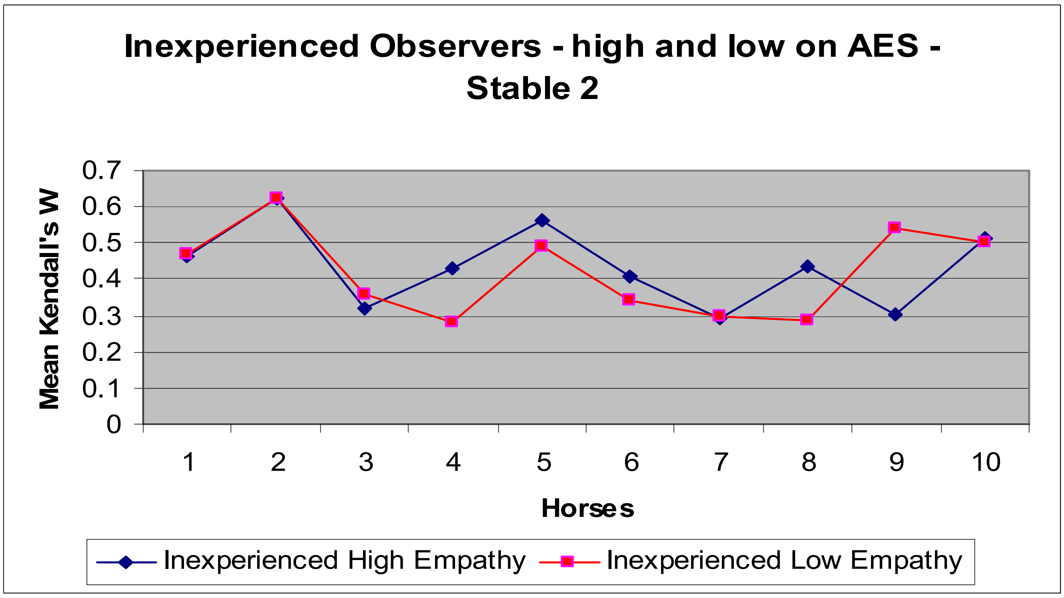


Figure 9.8

Kendall's W for inexperienced observers high and low on the empathy scale

The group mean Kendall's W values show, however, that both experienced and inexperienced observers lower on the AES also agreed less on horse personality scores (Figure 9.9).

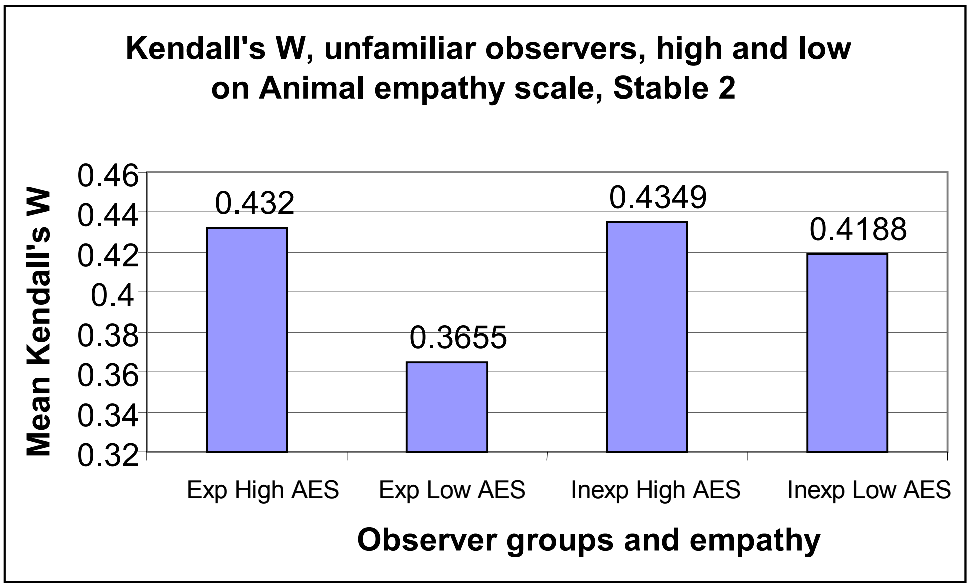


Figure 9.9

Kendall's W for high and low AES, Stable 2

To see if there is any significant difference in the degree of agreement between high and low empathy observers, we performed the t-test. There was no significant difference in the degree of agreement between the two inexperienced groups of observers; however, experienced observers high on AES agreed significantly more than less empathic experienced observers ($t=2.367$; $P<0.05$).

9.1.2.5. Degree of observers' empathy and horse personality scores

We treated each observer group separately, each with their own scoring list. There were not enough familiar observers to analyse their data in this part of the study. Therefore we used only data provided by the unfamiliar observers. The observers in both groups, experienced and inexperienced, were again divided into two groups: those who achieved animal empathy score of or greater than 147, and those whose score was lower than 147. We explored whether the degree of empathy influenced how the horses were positioned along the two dimensions. In other words, whether more empathic observers perceived horses to be higher or lower on the extroversion or the neuroticism dimension? Euclidean distance models are included in Appendix 30.

To investigate that, we correlated the horse personality scores generated by observers low and high on empathy. These correlations were done for experienced (Table 9.2) and inexperienced observers (Table 9.3).

Table 9.2

Correlation of horse personality scores provided by experienced observers high and low on empathy scale for Stables 1 and 2

Dimensions Spearman rho	**High Empathy Extroversion**	**High Empathy Neuroticism**
Low Empathy Extroversion	.917***	
Low Empathy Neuroticism		.411 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Correlations of horse personality score produced by participants high and low on the empathy scale was highly significant for the first dimension – extroversion (rho= 0.917, p<0.001), but not for the second dimension – neuroticism (rho=0.411, P>0.05) (Table 9.3).

Table 9.3

Correlation of horse personality scores provided by inexperienced observers high and low on empathy scale for Stables 1 and 2

Dimensions Spearman rho	**High Empathy Extroversion**	**High Empathy Neuroticism**
Low Empathy Extroversion	.871***	
Low Empathy Neuroticism		.822***

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Inexperienced observers that scored low or high on the AES, however, perceived and rated individual horses in a similar way. Scores generated by low and high empathic observers were significantly correlated for both dimensions, extroversion (rho=0.871, p<0.001) and neuroticism (rho=0.822, p<0.001).

9.1.2.6. Observers' personality and horse personality

The mean value in all observer groups for the neuroticism dimension (N) was 22 and for the extroversion dimension (E) it was 30. We split the observers into two groups: those who scored below the mean (low on N and E) and those who scored above the mean value (high on N and E). The two familiar groups only had four observers each so we could not do any further analysis on the data they provided.

There were only two experienced observers below the mean E score. Therefore in the experienced group of observers we calculated a separate MDS only for those who scored low or high on the neuroticism dimension of the NEO personality questionnaire.

Table 9.4
Correlation of horse personality scores provided by experienced observers high and low on neuroticism dimension for Stables 1 and 2

Dimensions Spearman rho	**Observers High on Neuroticism - Extroversion dimension (Horses)**	**Observers High on Neuroticism - Neuroticism dimension (Horses)**
Observers Low on Neuroticism - Extroversion dimension (Horses)	.9 ***	
Observers Low on Neuroticism - Neuroticism dimension (Horses)		.7 ***

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Horse personality scores produced by experienced observers who scored high on the N dimension and by observers low on N were significantly correlated. The spearman rho for the horse neuroticism dimension was 0.9 (p<0.001), and 0.7 (p<0.05) for the horse extroversion dimension (Table 9.4). Therefore, the observers' neuroticism dimension scores did not influence their assessment of horse personality.

In the inexperienced group we also split the observers into two groups, high and low on Extroversion and Neuroticism. There were six observers who scored high on the N dimension and 11 who scored low. In the same observer group there were seven observers who scored high on E and 10 who scored low. The MDS was generated separately for high and low on E and N. The MDS spatial configurations are included in Appendix 31.

Horse personality scores or spatial configurations provided by the observers low and high on E were then correlated (Table 9.5). The spearman rho shows that the observers high and low on E rated the horses from both stables in a similar way on both horse personality dimensions, extroversion (rho=0.851***, p<0.001) and neuroticism (rho=0.848***, p<0.001).

Table 9.5

Correlation of horse personality scores provided by inexperienced observers high and low on extroversion dimension, for Stables 1 and 2

Dimensions Spearman rho	**Observers High on Extroversion - Extroversion dimension (Horses)**	**Observers High on Extroversion - Neuroticism dimension (Horses)**
Observers Low on Extroversion - Extroversion dimension (Horses)	.851***	
Observers Low on Extroversion - Neuroticism dimension (Horses)		.848***

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Inexperienced observers who scored high on the neuroticism dimension also rated the horses in a similar way to the observers who scored low on the same dimension. Correlation of horse extroversion (rho=0.922, p<0.001) and neuroticism (rho=0.908, p<0.001) scores provided by the observers high or low on neuroticism was significant (Table 9.6). The spatial configuration of horses as provided by observers high or low on N is included in Appendix 31.

Table 9.6

Correlation of horse personality scores provided by inexperienced observers high and low on neuroticism dimension for Stables 1 and 2

Dimensions Spearman rho	**Observers High on Neuroticism - Extroversion dimension (Horses)**	**Observers High on Neuroticism - Neuroticism dimension (Horses)**
Observers Low on Neuroticism - Extroversion dimension (Horses)	.922 ***	
Observers Low on Neuroticism -		.908 ***

Neuroticism dimension (Horses)		

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

To sum up, the observers' E and N personality scores did not have a significant impact on how the observers rated the individual horses and how they construed the configuration of horses in a two dimensional space.

9.1.3. Synopsis

In this part of the first study we explored observers' individual differences and their influence on the degree of agreement and the spatial configuration of horses positioned along two personality dimensions. To assess observers' individual differences we used the Animal Empathy Scale and the NEO five factor personality questionnaire. The data was collected for all observer groups separately. However, there were only four observers in each familiar group so we did not submit the data they provided to inferential testing.

Secondly, both groups of unfamiliar observers, experienced and inexperienced, were divided into those who were high and those who were low on the Animal Empathy Scale. When observers rated horses from Stable 1, the results showed no consistent pattern in the association between AES and the degree of agreement. The unfamiliar observers with high AES scores agreed slightly more when they rated the horses from Stable 2. The degree of empathy also had no significant impact on the spatial configuration of horses positioned along the first dimension (extroversion); however, the configuration of horses along the second dimension (neuroticism) was affected by the empathy scores.

Finally, we were interested in finding out if the observers' personality influences their assessment of horses. We divided the unfamiliar observers high and low on E and N according to the personality scores collected with the NEO personality questionnaire. We then generated two dimensional spatial configurations of horses, using MDS. These configurations were generated separately for the observers low and those high on E and N. All correlations of the results provided by observers high and low on E and N were significant, which means that observers' personality had no significant effect on how they positioned the horses in that two dimensional spatial configuration.

9.2. STUDY 2

In this study we will first present information about the results from observers' personality and emotional intelligence (EI) questionnaires. Then we will look at the relationship between EI scores and the degree of agreement. Finally, we will correlate the horse personality scores provided by the observers with low and high scores on EI, extroversion and neuroticism.

9.2.1. Methods

In the second study two groups of observers, familiar and unfamiliar, rated three groups of horses from three different stables. Familiar observers rated only their own horses i.e. staff from Stable 1 assessed only their own horses and staff from Stable 2 assessed only their own horses. In total there were 34 observers and 38 horses. Familiar observers were staff who had worked with the horses over a long period of time. Unfamiliar observers did not know the horses participating in this study and may or may not have had experience with horses. This experiment includes only one unfamiliar group as there were no consistent differences between experienced and inexperienced observers in the first study. Unfamiliar observers assessed the horses from Stables 1 and 2 (see Figure 7.2). Familiar observers rated the horses based on their previous experience with the horses, unfamiliar on the basis of video clips of each individual horse (see section 8.1.3.).

The observers completed the NEO Five Factor Personality questionnaire (Costa & McCrae, 1992) and an emotional intelligence questionnaire (Schutte et al., 1998), which is a 33 item questionnaire (a copy of the questionnaire is included in Appendix 32).

In this part of the Study 2 we investigated whether the observers' individual differences, emotional intelligence and personality, influence their assessment of horse personality. Our hypothesis was that observers' individual differences would affect their horse personality scores.

9.2.2. Results

We collected the observers' background information: age, personality and emotional intelligence. The overall mean age of the observers was 25.1 years. Observers from Stable 1 had the highest mean age (31.4), while unfamiliar observers had the lowest mean age of 20.5. The results from the NEO five factor personality questionnaire (Costa & McCrae, 1992) show that the observers from Stable 1 had the lowest scores for agreeableness, but scored highest on conscientiousness and extroversion. Observers from Stable 3 had the highest scores for openness (the five factor personality model was reviewed and discussed in section 3.2.2.4).

The emotional intelligence questionnaire (Schutte et al., 1998) showed that the unfamiliar observers were not only the youngest, but were also the highest on the EI scale and neuroticism, and the lowest on conscientiousness; and that they were the most agreeable (Table 9.7).

Table 9.7
Observers' Individual differences

	Age (mean)	Emotional Intelligence (mean)	Agreeableness (mean)	Conscien-tiousness (mean)	Extrover-sion (mean)	Neuroti-cism (mean)	Openness (mean)
Stable 1 Familiar	31.4	112.71	26.2	37.4	35	19.7	27
Stable 2 Familiar	24.2	117.13	32.6	36.7	30.8	19	24
Stable 3 Familiar	24.6	117.89	30	34	28.6	20.3	29.4
Unfamiliar observers	20.5	118.70	31	30.8	28.4	25.5	27.8
Overall mean	25.1	116.6	29.9	34.7	30.7	21.1	27
Standard deviation	4.5	2.6	2.7	2.9	3	2.9	2.2

9.2.2.1. Emotional intelligence scale

The Emotional intelligence scale (Schutte et al., 1998) is a 33-item self-report measure based on Salovey and Mayer's (1990) early work. Emotional intelligence is defined as one's ability to recognize and monitor their own emotions and the emotions of others. This includes the ability to discriminate among different emotions and to communicate them with others (Schutte et al., 1998).

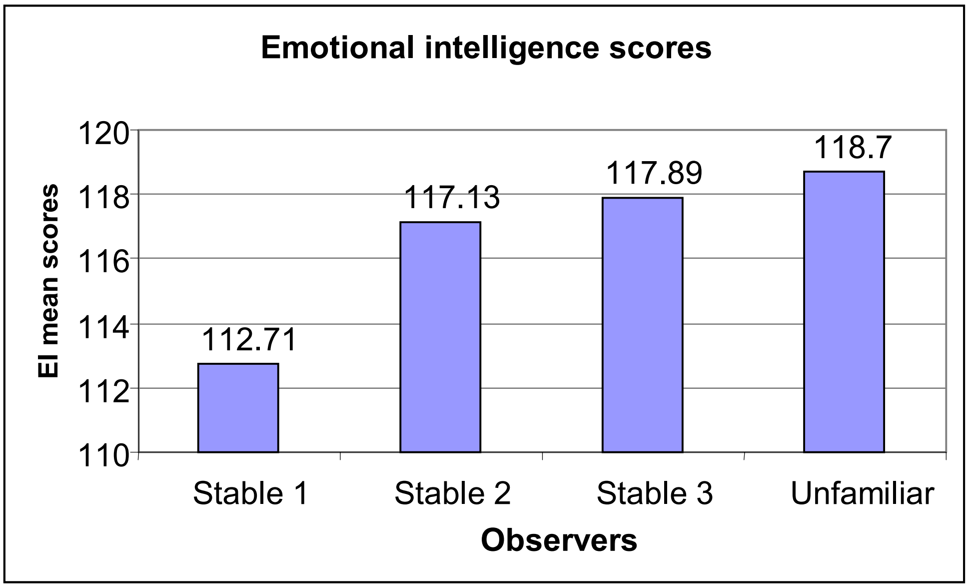


Figure 9.10
Emotional intelligence scores for four observer groups

The mean values for different observers indicate that EI scores are lower for the older observer group, while the unfamiliar observers, who were the youngest, had a higher mean EI value (Figure 9.10).

9.2.2.2. EI and observers' degree of agreement

The main question in this section was whether observers' ability to monitor others' emotions had any influence on their degree of agreement when assessing the horses.

The overall mean for EI was 117. We divided the observers in each observer group into two groups; those who scored high and those who scored low on EI. We calculated the Kendall's coefficient of concordance W for observers high and low on EI. There was no consistent pattern in the degree of agreement between observers high and those low on EI (Figure 9.11).

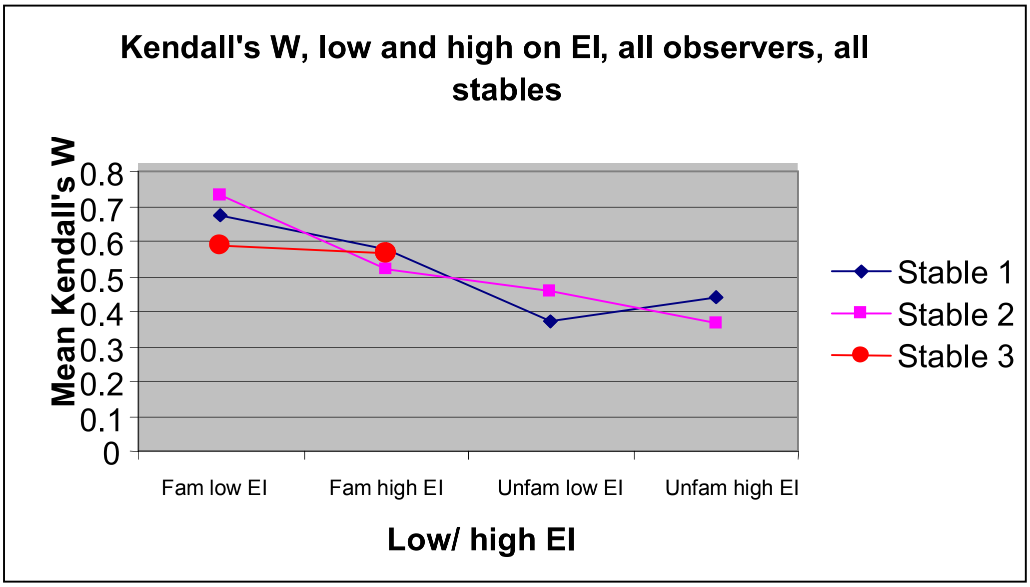


Figure 9.11

Kendall's W for high and low EI scores, all stables and all observer groups

In Stable 1, familiar observers high on EI had a significantly lower degree of agreement than observers low on EI ($t=2.646$, $p<0.05$). At the same time unfamiliar observers high on EI had a significantly higher degree of agreement than those low on EI ($t=2.449$, $p<0.05$). In Stable 2 the familiar observers with high EI also had lower degree of agreement than the observers with low EI ($t=6.56$. $p<0.001$). However, there was no difference in the degree of agreement between high and low EI in Stable 3 (a table of Kendall's W values for all stables, with observers high and low on EI is included in Appendix 33).

9.2.2.3. *EI and horse personality dimensions*

We were first interested in whether high or low scores on the EI questionnaire influenced observers' assessment of horse personality. The EI mean of 117 was a cut off point here as well. We calculated spatial configuration of horses separately for the observers low and those high on the EI questionnaire. We then correlated the two sets of values to see whether the observers low on EI mapped horses in a similar way to the observers high on the EI scale.

In the familiar group from Stable 1 three observers scored low on EI and four scored high. The familiar observers from Stable 1 positioned the horses in a significantly similar way along the first dimension, which we labelled agreeableness (rho=0.984, p<0.001), but not along the second dimension (Table 9.11), which we labelled extroversion.

Table 9.11

Correlation of horse personality scores generated by familiar observers high and low on Emotional intelligence scale –Stable 1

Dimensions **Spearman rho**	**High EI** **Dimension 1 A**	**High EI** **Dimension 2 E**
Low EI **Dimension 1 A**	.984 ***	
Low EI **Dimension 2 E**		.401 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

In Stable 2 there were three familiar observers who scored low on EI and five who scored high. The correlation coefficient shows that the observers located the horses in a similar way on both dimensions (Table 9.12)

Table 9.12

Correlation of horse personality scores generated by familiar observers high and low on Emotional intelligence scale –Stable 2

Dimensions Spearman rho	**High EI Dimension 1 N**	**High EI Dimension 2 E**
Low EI Dimension 1 N	.863 ***	
Low EI Dimension 2 E		.670 *

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

In Stable 3 there were three familiar observers who scored low on EI and six who scored high. They agreed on positioning the horses along the neuroticism dimension (rho=0.622, p<0.05), but not on the agreeableness dimension (Table 9.13).

Table 9.13

Correlation of horse personality scores generated by familiar observers high and low on the Emotional Intelligence scale – Stable 3

Dimensions Spearman rho	**High EI Dimension 1 N**	**High EI Dimension 2 A**
Low EI Dimension 1 N	.622 *	
Low EI Dimension 2 A		.084 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

In the unfamiliar group five scored low on EI and five scored high. As the unfamiliar observers rated both groups of horses with the same list of adjectives, we correlated the scores for all horses provided by high and low EI observers. Both dimensions were significantly correlated, so the, EI did not affect rating of the horses (Table 9.14). Spatial configurations for all observer groups, low or high on EI, are included in Appendix 34 and Appendix 35.

Table 9.14

Correlation of horse personality scores generated by unfamiliar observers high and low on the Emotional Intelligence scale – Stables 1 and 2

Dimensions **Spearman rho**	**High EI** **Dimension 1 (E-Stable2; N-Stable1)**	**High EI** **Dimension 2 (E/N-Stable2; E-Stable1)**
Low EI **Dimension 1**	.603 ***	
Low EI **Dimension 2**		.564 **

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

In brief, the emotional intelligence scores of the observers did not have a significant effect on the rating of horses on the first dimension. However, the familiar observers with different EI scores from the first and the third stable positioned the horses differently along the second horse personality dimension.

9.2.2.4. Observers' personality (NEO) and positioning of the horses

In this part we positioned all observers either low or high along two personality dimensions: neuroticism (N) and extroversion (E). The scores were obtained using the NEO personality questionnaire (Costa & McCrae, 1992).

9.2.2.4.1. The neuroticism dimension

The overall mean score for the neuroticism dimension was 21. This was the cut off point for positioning the observers either high or low on N. There were different numbers of observers high and low on N at each stable. At the first stable three observers scored low on N and four scored high. Horse personality scores provided by the observers high and low on N were significantly associated for the first dimension (rho=0.885, $p<0.05$), but not for the second (Table 9.15).

Table 9.15

Correlation of horse personality scores generated by familiar observers high and low on Neuroticism dimension – Stable 1

Dimensions **Spearman rho**	**High Neuroticism** **Dimension 1 A**	**High Neuroticism** **Dimension 2 E**
Low Neuroticism **Dimension 1 A**	.885 ***	
Low Neuroticism **Dimension 2 E**		.374 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

In Stable 2 the familiar observers both high and low on N also agreed on the first dimension, but not on the second (Table 9.16). Here we had six observers who scored low on N and two who scored high.

Table 9.16

Correlation of horse personality scores generated by the familiar observers high and low on the Neuroticism dimension – Stable 2

Dimensions **Spearman rho**	**High Neuroticism** **Dimension 1 N**	**High Neuroticism** **Dimension 2 E**
Low Neuroticism **Dimension 1 N**	.577 *	
Low Neuroticism **Dimension 2 E**		.401 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

A similar pattern is present in Stable 3. There were five observers low and four high on N dimension. The two groups of observers again positioned the horses in a similar way along the first horse personality dimension (rho=0.678, $p<0.05$), but not along the second (Table 9.17).

Table 9.17

Correlation of horse personality scores generated by familiar observers high and low on Neuroticism dimension – Stable 3

Dimensions Spearman rho	**High Neuroticism Dimension 1 N**	**High Neuroticism Dimension 2 A**
Low Neuroticism Dimension 1 N	.678 *	
Low Neuroticism Dimension 2 A		.105 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Unfamiliar observers scored the horses from stables 1 and 2 on the same list of constructs, so we merged the data from both stables and calculated separate spatial configurations of horses for the observers high and the observers low on N. There were three observers low on N and seven high. Observers' scores on the N dimension did not influence how they mapped the horses along the two horse personality dimensions (Table 9.18).

Table 9.18

Correlation of horse personality scores generated by unfamiliar observers high and low on Neuroticism dimension – Stables 1 and 2

Dimensions Spearman rho	**High Neuroticism Dimension 1**	**High Neuroticism Dimension 2**
Low Neuroticism Dimension 1	.583 **	
Low Neuroticism Dimension 2		-.553 **

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

To sum up, whether observers scored high or low on the neuroticism dimension did not affect how they mapped the horses along the first dimension. It also did not affect how the unfamiliar observers positioned the horses along dimension two. However, there is no significant correlation between the familiar observers high and low N for dimension two. The Euclidian distance models for each observer group high or low on N, are included in Appendix 36.

9.2.2.4.2. The extroversion dimension

The overall mean score for the extroversion dimension was 31, which was the cut off point when we divided the observers according to whether they scored low or high on E. In Stable 1 all familiar observers but one were high on E, so we did not perform MDS on their data.

In Stable 2 three familiar observers were low and five were high on E. The observers' extroversion did not significantly affect their configuration of horses along the first dimension (rho=0.901, $p<0.001$). However, the way they positioned the horses along the second dimension was not significantly correlated (Table 9.19).

Table 9.19
Correlation of horse personality scores generated by familiar observers high and low on the Extroversion dimension – Stable 2

Dimensions Spearman rho	**High Extroversion Dimension 1 N**	**High Extroversion Dimension 2 E**
Low Extroversion Dimension 1 N	.901 ***	
Low Extroversion Dimension 2 E		.412 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

In Stable 3 six familiar observers were low on E and three were high. Again the differences in their extroversion scores did not influence correlation of the first dimension. However, the horse personality scores they provided for the second dimension were not significantly related (Table 9.20).

Table 9.20

Correlation of horse personality scores generated by familiar observers high and low on Extroversion dimension – Stable 3

Dimensions Spearman rho	**High Extroversion Dimension 1 N**	**High Extroversion Dimension 2 A**
Low Extroversion Dimension 1 N	.650 *	
Low Extroversion Dimension 2 A		.154 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Ten unfamiliar observers, seven of whom were high on E, also mapped the horses along two horse personality dimensions. Similarly to EI and N, E did not affect the scoring as the scores for both dimensions, provided by the observers high and low on E, were significantly correlated (Table 9.21).

Table 9.21

Correlation of horse personality scores generated by unfamiliar observers high and low on Neuroticism dimension –Stables 1 and 2

Dimensions Spearman rho	**High Extroversion Dimension 1**	**High Extroversion Dimension 2**
Low Extroversion Dimension 1	.75 ***	
Low Extroversion Dimension 2		.706 ***

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

On the whole, the observers' extroversion scores did not affect the first horse personality dimension, although it may have affected the second dimension as the spatial configurations of horses along dimension 2, provided by the familiar observers high and low on E, were not significantly correlated. The extroversion of the unfamiliar observers did not have a significant effect on their scoring as the horses were mapped in a similar way along the two dimensions. The spatial configurations for all observer groups low and high on E are included in Appendix 37.

9.3. Synopsis

In this chapter we have reported how observers' individual differences, their personality, empathy and emotional intelligence influenced their assessment of horses. In both studies we used familiar and unfamiliar observers. In the first study we further divided the unfamiliar group into those experienced and those inexperienced with horses. In the second study we only had familiar and unfamiliar observers.

We were expecting observers' empathy, personality and emotional intelligence to influence their rating of horses. However, there was no indication that the animal empathy scores or emotional intelligence scores influenced either the degree of agreement or the mapping of the horses along two dimensions. We also divided the observers accordingly to whether they were low or high on neuroticism and extroversion and separately calculated their spatial configuration of horses in a two dimensional space. Observers' extroversion and neuroticism had no effect on positioning the horses along the first dimension. Observers' personality also had no impact on how the horses were located along the second dimension by the unfamiliar observers. However, in most familiar observer groups the observers high on neuroticism and extroversion positioned horses differently along the second dimension from the observers low on extroversion and neuroticism.

In conclusion, the results show that in most cases the observers' personality, empathy and emotional intelligence does not influence their assessment of horse personality.

CHAPTER 10

Discussion, Achievements, Limitations and Further Recommendations

10.1. Introduction

We started the project by formulating three questions:

1) How reliable is the Repertory grid technique as a method for assessing horse personality?
2) What are the horse personality dimensions?
3) Do observers' individual differences influence their assessment?

The results show high reliability of the repertory grid technique with significant agreement in all observer groups. The personalities of 56 out of a total 59 our horses were agreed on by 78 observers. They observers were grouped in eight groups, and the results refer to *within* group agreement. This indicates that the repertory grid technique incorporates a low level of measurement error and high reliability. Familiar observers manifested significantly higher degree of agreement than unfamiliar observers. The results also show that the constructs generated and defined by observers were used highly consistently. There were 179 constructs in total of which 159 were agreed on and therefore applied consistently and systematically by the raters. This result contradicts a previous study, which applied the same method of data collection and reported that only two out of 20 of descriptors generated by the observers were used consistently (Mills, 1998).

We discovered three horse personality dimensions: extroversion, neuroticism and agreeableness with only familiar observers generating the agreeableness dimension. These results are consistent with the findings in most animal personality studies (Gosling, 2001).

Finally, horse personality scores provided by observers high and low on empathy, emotional intelligence, neuroticism and extroversion were similar. Twenty out of 26 of horse personality dimensions provided by observers with different personality profiles were significantly correlated in both studies.

Our three main questions were further specified as eight hypotheses, introduced in Chapter 4:

1) Both familiar and unfamiliar observers will provide similar personality descriptors for horses.
2) Familiar observers will have a higher degree of agreement when rating the horses.
3) Observers will rate more reliably on the adjectives they provide themselves.
4) Observers will use the constructs consistently.
5) Horse personality dimensions will be in line with the five factor personality model.
6) Familiar and unfamiliar raters will agree on their horse personality scores.
7) There will be an association between horse personality and performance.
8) Observers' personality, empathy and emotional intelligence will not affect their rating of the horses.

Hypothesis one refers to the elicitation part of our project. Hypotheses two to four refer to the reliability of repertory grid technique as a method for assessing horse personality. Hypotheses five to seven refer to our second question, and hypothesis eight elaborates our question three. In the following sections of this chapter we will first discuss the main results according to the three main questions and more specific hypotheses, then address the achievements and the limitations of the project, and finally make further recommendations.

10.2. Horse Personality Descriptors

Our first hypothesis was that familiar and unfamiliar observers will provide similar personality descriptors for horses. In the first study each observer group generated their own list of constructs, while in the second study there were only two scoring lists, one for the familiar and one for the unfamiliar observers. Some descriptors were present in all scoring lists, e.g. "friendly", "shy", "interested" and "gentle". "Friendly" was paired with "unfriendly" by all observer groups in the two studies; however, "affectionate" was paired with "aloof" in the first study and "aggressive" in the second study. According to Kelly (1955) words are only labels and we need to discover the meaning behind them. The observers were interviewed individually after the elicitation and asked to elaborate the constructs. Often slightly different words were used for the same meaning. Descriptors generated only by the familiar observers were "obedient-disobedient", "easy to lead", "easy to handle" and "willing to work". These were associated with "agreeableness", which was not present among the constructs of the unfamiliar observer groups. Provided scoring lists tend not to include constructs associated with agreeableness (see Sian-Lloyd, 2006, in press). Lists of descriptors reported by other researchers are versatile. Worth-Estes (1952) composed a 31-item scoring list generated by familiar observers. The descriptors included were "follower", "difficult to catch", "very affectionate" and "nervous". Anderson's et al. (1999) list, generated by riding instructors, includes "cooperative-uncooperative", "calm-nervous" and "gentle-rough". Momozawa et al. (2003) also include similar adjectives: e.g. "friendly", "nervous" and "stubborn". Studies that have adapted descriptors (Sian-Lloyd et al, 2006, in press) or questions (Morris et al.,

2002) from other studies (often from Stevenson-Hinde and Zunz, 1978) do not include descriptors that are more specific for horses and important for the observers, e.g. "easy to lead" and "easy to handle". To sum up, in this project we have created six different scoring lists, generated by both familiar and unfamiliar observer groups. They are all based on horse behaviour and use similar descriptors. Our lists of constructs are most similar to the one presented by Worth-Estes (1952).

10.3. Reliability of the Repertory Grid Technique

As discussed in Chapter 2 one way to determine the reliability of a method is to calculate inter-observer agreement: the higher the agreement the more reliable the method. Significant agreement is determined by the question as to whether we can get the same results more than once or from more than one judge. Significant degree of agreement between judges means that there is a low level of measurement error.

Reliability therefore refers to dependability, stability and consistency of behavioural measurements or judgements. In most farm animal personality studies reliability is measured by the consistency of individual behaviours. Lawrence et al. (1991) investigated the consistency of "temperamental characteristics" in pigs. They concluded that they are not consistent over time and situations. These results, however, should be further investigated as test-retest correlation shows that personality traits can change over time (Matthews & Deary, 1998). In our study we follow the work of Stevenson-Hinde and Zunz (1978), French (1993) and Wemelsfelder et al. (2000), where the main question is how stable and consistent is observers' assessment of animals' behaviour.

In this section we will discuss our hypotheses two, three and four: that familiar observers have a higher degree of agreement when rating the horses; that observers rate more reliably on the adjectives they provide themselves; and that observers use the constructs consistently. In both studies we calculated the degree of agreement for horses and the constructs. We were interested in the reliability of horse personality rating and in the consistency of the use of personality descriptors. The results showed a significant degree of agreement for both horses and constructs. We further investigated whether familiarity with horses and with words influenced the degree of agreement.

Overall degree of agreement was calculated for each individual horse and we found that the observers in all observer groups agreed on 56 out of 59 horses. These results refer to agreement within each observer group, not between them. Only on the personality of three horses did observers not agree significantly: in Study 1 the familiar observers did not agree on Sam and Slotty; and in Study 2 the unfamiliar observers did not agree on Tia. Their personality profiles show that the observers may have had difficulties deciding about their consistent behaviour and rated them with scores of around three (on the scale from one to five). On the other hand horses with a higher degree of agreement, e.g. Ginger and Appley, had more extreme mean personality scores. Therefore with more extreme behaviour that is easier to observe, they may represent so called "good targets" (Funder & Colvin, 1988; Funder, 1997; Sian-Lloyd et al., 2006 in press).

10.3.1. Familiarity with horses

We were expecting the familiar observers to have a higher degree of agreement (Funder 1997). The results from Study 1 show that the familiar observers from the first stable had a significantly higher Kendall's W than inexperienced observers ($t=2.78$, $p<0.05$). The familiar observers from the second stable had a significantly higher degree of agreement than both experienced ($t=3.79$, $p<0.01$) and inexperienced observers ($t=3.14$, $p<0.05$). In Study 2 familiar observers from Stable 1 agreed significantly more than unfamiliar observers ($t=9.5$, $p<0.001$). The results were similar for Stable 2: the difference in degree of agreement between familiar and unfamiliar observers was significant ($t=4.7$, $p<0.001$).

Our results show that familiarity with animals increases inter-rater agreement significantly and that the assessment by familiar observers is more reliable. This confirms the arguments put forward by previous researchers (Martau et al., 1985; Lyons, 1989) that familiarity with animals will increase the accuracy of the assessment. These results are also in line with the research in human psychology (Funder & Colvin, 1988), arguing that increased familiarity with targets will also increase correlation between judges and therefore make the assessment more reliable.

10.3.2. Familiarity with words

We predicted that the observers would agree more on the constructs they generated themselves than on constructs that were partially provided to them. In Study 1 the familiar observers had a significantly higher degree of agreement when using elicited constructs for Stable 1 ($t=2.24$, $p=0.05$). However, when inexperienced observers rated the horse personality on elicited constructs, they agreed more for horses from Stable 1 ($t=4.89$, $p<0.001$) and Stable 2 ($t=5.2$, $p<0.001$).

The results from Study 1 indicate that familiarity with words increases the degree of agreement and reliability of the assessment. This finding supports Adams-Weber (1970), Dutton et al. (1997) and Wemelsfelder et al. (2000), who argue that reliability of the rating is increased when observers are given the opportunity to be involved in the assessment from the beginning and to integrate their perception of animals. Most animal personality studies, however, use provided scoring lists, which may judging by our results, affect the reliability of the assessment.

The results from the second study, however, did not confirm the results from Study 1. Contrary to our expectations unfamiliar observers had a significantly higher degree of agreement when rating the horses on the constructs that were partially provided to them, both in Stable 1 ($t=2.803$, $p<0.05$) and Stable 2 ($t=2.507$, $p<0.05$). There were no significant differences in the familiar observer groups. The design for the second study was slightly altered and there may be different reasons for the inconsistent results: first, the scoring was done in groups; secondly, the experimenter spent less time with individual observers during the interviews to elaborate their constructs; thirdly, the uniform scoring lists were further uniformed as there was only one scoring list for familiar and one for unfamiliar observers; and finally, the unfamiliar raters in Study 2 were less motivated than the observers in Study 1. The dropout rate was almost 50 percent in comparison to none in the first study.

Another important point is that all constructs were essentially elicited by the observers themselves. The constructs we labelled as partially provided were just generated by a lower number of participants. Therefore the distinction between elicited and "partially provided" may not be very clear, which could be reflected in the results. Fransella and Bannister (2004) argue that the provision of constructs does not really affect the accuracy of scoring as long as they are supplied to a population that can apply them. As mentioned above, all our constructs were elicited based exclusively on horse behaviour.

10.3.3. Observers' consistency in applying personality constructs

Our fourth hypothesis was that the observers would use the constructs consistently, which would also indicate how reliable the method is.

In Study 1 there were four scoring lists, with a total of 117 constructs. One-hundred and two constructs in total had a significant Kendall's W and therefore were used consistently. Constructs with the highest degree of agreement and therefore applied most consistently were "laid back – worrier" (Stable 1, familiar observers, $W=0.815$, $p<0.001$), "affectionate – aloof" (Stable 2 familiar observers, $W=0.737$, $p<0.001$) "relaxed – tense" (experienced observers, Stable 1, $W=0.599$, $p<0.001$), "quiet – energetic" (experienced observers, Stable 2, $W=0.714$, $p<0.001$), "sociable – unsociable" (inexperienced observers, Stable 1, $W=0.534$, $p<0.001$) and "docile – spirited" (inexperienced observers, Stable 2, $W=0.766$, $p<0.001$).

In Study 2 we had only two scoring lists with 62 constructs in total. Fifty-seven constructs were used consistently. The most consistent constructs were "bargy-quiet" (familiar, Stable 1, $W=0.677$, $p<0.001$), "mature-immature" (familiar, Stable 2, $W=0.699$, $p<0.001$), "confident-timid" (familiar, Stable 3, $W=0.630$, $p<0.001$), "energetic-lazy" (unfamiliar observers, Stable 1, $W=0.463$, $p<0.001$) and "affectionate-cold" (unfamiliar observers, Stable 2, $W=0.556$, $p<0.001$).

The constructs used most consistently are all different: some are associated with neuroticism and some with extroversion.

To sum up, within their observer groups the observers agreed on the personality of 56 out of 59 horses. Familiarity with the individual horses increased the degree of agreement significantly. Animal personality studies often use experts to generate the constructs and rate animals on them. Our results showed that having general information about the species did not significantly increase reliability; however, having information about individual animals did. We can therefore conclude that familiar observers provided a significantly more reliable rating of horse personality. We predicted that familiarity with the constructs would also increase observers' agreement and reliability of the method. The results from the inexperienced group from Study 1 supported this; however none of the other observer groups confirmed it. The observers applied 159 out of 179 constructs consistently and systematically. Our findings do not support the results from a previous study (Mills, 1998), which reported inconsistent use of individually generated and defined constructs and consequently recommended "objectively predefined adjectives" to reliably assess horse personality.

10.4. Horse personality dimensions

In this section we will discuss the following three hypotheses: 5) that horse personality dimensions will be in line with the five factor personality model; 6) that familiar and unfamiliar raters will agree on their horse personality scores; 7) and that there will be an association between horse personality and performance.

The results showed three personality dimensions in horses. We labelled them according to the Five factor personality model (Costa & McCrae, 1992): extroversion, neuroticism and agreeableness. Not all these dimensions were clearly defined in all observer groups. Agreeableness was absent in the assessment of unfamiliar groups, regardless of their previous experiences with horses. There were not many significant correlations between the observer groups. The agreement between the observers *within* each group was high; however, mostly non-significant correlations of personality scores *between* the observer groups show that the observers from different groups positioned the horses differently along the two dimensions. Only four out of 12 correlations were significant in Study 1 and two out of six in Study 2.

We also presented horse personality profiles. In the first study, in relation to the highest and the lowest degrees of agreement; and in the second study, in relation to horses' performance scores.

10.4.1. Identification of horse personality dimensions

In the first study, in Stable 1 all observer groups generated two dimensions, labelled as neuroticism and extroversion. In Stable 2 the familiar observers came up with neuroticism and extroversion/agreeableness; the latter included constructs that could be associated with both dimensions. The two unfamiliar groups produced neuroticism and extroversion.

In the second study in Stable 1 the familiar observers produced agreeableness and extroversion, the unfamiliar neuroticism and extroversion. In the unfamiliar group there was an overlap of constructs significantly associated with both dimensions. In Stable 2, the familiar observers produced neuroticism and extroversion. The unfamiliar observers had a strong first dimension, extroversion, and an unclear second dimension which we labelled extroversion/neuroticism. In Stable 3 the familiar observers produced neuroticism and agreeableness.

These three dimensions seemed to be consistently present in all our horses. The findings are in line with Le Scolan et al. (1997), who report "nervousness" and "sociability" based on the constructs produced by riding teachers. Morris et al. (2002) and Sian-Lloyd et al. (2006, in press) also report high reliability for three dimensions: agreeableness, neuroticism and extroversion. Our findings are therefore consistent with previous research on horses and other species (Gosling, 2001). Neuroticism and extroversion are two of the three super-factors of Eysenck's biologically based model of personality model of human personality with strong biological roots (Eysenck, 1952). The third super-factor is psychoticism, which could be associated with constructs that form our agreeableness dimension. Consistency of our findings with previous research on animal and horse personality (Gosling, 2001; Morris et al., 2002) indicates the validity of the RGT as a method for assessing horse personality.

Labelling of the dimensions was arbitrary for two reasons: firstly, the second dimension was in some cases not clearly defined, particularly in Study 2; and secondly, the same constructs could be interpreted as relating to facets of more than one of the Big Five personality dimensions. Our decisions were based on the facets that form the five different personality dimensions in the Five Factor Model of personality (Costa & McCrae, 1992; Gosling, 2001). It is important, however, to emphasise that this labelling endorses reductive assessment and was done purely for comparison with previous animal personality studies and with meta-analysis by Gosling (2001).

10.4.2. Differences between familiar and unfamiliar observers

We predicted that familiar and unfamiliar observers would agree on their assessment of horse personality and how they mapped the horses in two dimensional space. The results, however, do not support our hypothesis.

In Study 1 there was significant correlation between familiar and inexperienced observers on the first dimension, neuroticism (rho=0.818, $p<0.01$). There was also significant correlation between experienced and inexperienced observers on dimension 2, extroversion rho=0.952, $p<0.001$). In Stable 2 there was no correlation between familiar and unfamiliar observer groups, although, the experienced and inexperienced observers mapped the horses similarly on both dimensions, neuroticism (rho=0.685, $p<0.05$) and extroversion (rho=0.661, $p<0.05$). In Study 2 there was a significant correlation between familiar and unfamiliar observers from Stable 1 for dimension 1 (rho=0.55, $p<0.05$) and a trend towards significant correlation was shown for dimension 2 (rho=0.522, p =0.067). However, there was no significant correlation between familiar and unfamiliar observers in Stable 2.

There may be several reasons for that. Firstly, the familiar observers knew the horses well, while the unfamiliar did not. We have already discussed the differences in agreement between the familiar and unfamiliar observers. Secondly, the experimental design for familiar and unfamiliar observers was different. While the familiar observers rated their horses based on their previous experience with them (Dutton et al., 1997), the unfamiliar observer groups watched a short video of individual horses in a human interaction test, and were therefore observing behavioural styles (Wemelsfelder et al., 2000), which are intersubjective operational definitions of emotions (Crossley, 1996). Finally, it is quite possible that we were measuring two different phenomena: personality with familiar observers and emotions with unfamiliar observers. A way to test that would be to show the videos to the familiar observers and ask them to rate the horses again. After the first study we aimed to increase the versatility of information available to the unfamiliar observers on the video. So we added a standardised novel object test. The results were slightly different - in Stable 1 the scores produced by the familiar and unfamiliar observers for the first dimension were significantly correlated. There was also a trend towards correlation for the second dimension; however there was no correlation or trends towards it for Stable 2.

Since the findings are inconsistent, we cannot conclude whether the inconsistency in personality scores between familiar and unfamiliar observer groups is due to observers' familiarity with horses or due to the differences in experimental design. However, if we speculate, it is possible that the results do not reflect on the validity of the RGT, but simply mean that we measured two different things, personality with familiar observers and emotions with unfamiliar ones.

10.5. Personality profiles of individual horses

In both studies we presented line graphs for several horses with their mean Kendall's W values for each construct. None of the previous horse studies (e.g. French, 1993; Mills, 1998; Morris et al., 2002) presented detailed information about individual horses. However, pig personality profiles have been previously calculated by Wemelsfelder et al. (2000).

In the first study our aim was to show personality profiles for horses which the familiar observers agreed on the most and the least. We chose to present the profiles made by familiar observers as they knew the horses well. Sam and Slotty were the two horses with the lowest degree of agreement. For Sam, most scores were around three on the five point rating scale, with few extreme scores. Sam was assessed as a worrier, who was unreliable and unwilling to work. Slotty was rated as slightly aloof and sensitive. On the other hand Ginger and Apply both received the highest raters' agreement. Their scores were more extreme and they were perceived as more agreeable and easy to work with.

Based on these findings we asked the familiar observers in Study 2 to rate their horses on a single item performance scale, and we subsequently calculated personality profiles for horses which were rated as the best and the worst performers. In Stable 1 Zak and Gunther received the lowest performance score. Zak was rated as difficult to work with, however Gunther's profile was more similar to Barny's and Blue's – two horses that were the best performers. Again they were perceived as easy to work with and nice. In stable two Lil and Kurt were low performers. They were both rated as inexperienced and immature. Rocky and Jazz, the best performers, were again easy to handle and work with. In Stable 3 Sophie and Fleur performed badly. Fleur was immature and inexperienced, while Sophie was excitable and highly strung. Bruno was nice, patient and interested and Jester more bossy and independent.

Therefore the horses the observers agreed on the most had more similar personality profiles to horses with high performance. The two horses with a low agreement were, similarly to horses with low performance rate, disagreeable, nervous and inexperienced. Most previous studies (e.g. Anderson et al., 1999; Visser et al., 2003b) found no consistent relationship between the results from personality ratings and performance. However, our exploration indicates that there may be a relationship between horses' performance and their agreeableness traits.

10.6. Observers' individual differences

Our last hypothesis was tied up with the last question we formulated at the beginning of this project: Will observers' personality, empathy and emotional intelligence affect their rating of the horses?

The results show that observers' individual differences do not consistently influence their assessment of horses. Seventy-seven percent of horse personality ratings provided by observers with high and low on empathy, emotional intelligence and personality scales, were significantly correlated. Very often assigning animal personality is dismissed as anthropomorphic and a projection of observers' personality onto animals. We have already tackled the question of anthropomorphism in Chapter 5. In the last chapter, however, we empirically addressed the question whether observers' individual differences significantly affect their rating scores. No animal personality study has ever attempted to do that; therefore these results may be a significant contribution to the discussion on self projection and anthropomorphism.

In Study 1 we calculated Kendall's W separately for observers high and low on the Animal Empathy Scale (AES) (Paul, 2000). Inexperienced observers were the only group where observers with higher empathy scores agreed significantly more on rating the horses ($t=2.367$; $p<0.05$). We performed separate multi dimensional scaling (MDS) for unfamiliar observers high and low on the AES and correlated the scores. Correlations were significant for all but one dimension. Therefore in all cases but one the results from the AES did not have an effect on horse personality scores. We did the same with observers low and high on extroversion and neuroticism. The horse personality scores provided by the observers high and low on those two dimensions were significantly correlated. Therefore the observers' neuroticism and extroversion scores had no significant effect on their ratings.

In Study 2 the horse personality scores provided by observers high and low on emotional intelligence (EI) were correlated. The results show that only the second dimension provided by the familiar observers from Stables 1 and 3 were not significantly correlated. Therefore apart from in these two cases the EI has no affect observers' assessment either. We then further split the observers into high and low on two personality dimensions, neuroticism and extroversion. Familiar observers with high or low scores on E and N mapped the horses in the same way along the first dimension, but not the second. However, in the unfamiliar group the observers' personality had no significant effect on horse personality scores.

To sum up, only six out of 26 correlations of horse personality scores provided separately by the observers low or high on individual differences scales were not significantly correlated. This is not necessarily due to observers’ individual differences. We already mentioned that the second dimension in Study 2 was not clearly defined, which may be reflected in the results. The number of observers scoring low or high on their personality, empathy or emotional intelligence was not equally spread, which might have affected the mapping of the horses as well. Furthermore, the finding that 20 out of 26 correlations of horse personality scores generated by the observers who differ in their personality, empathy and emotional intelligence, were significant indicates that different observers still rated the horses in a similar way. We would expect different results, if the assessment of horse personality with individually generated and defined adjectives was only a self-projection.

10.7. Achievements of the project

There are three important findings in this project. The first is the reliability of the Repertory grid technique as a method for assessment of horse personality. Within each observer group the observers agreed on the personality of 56 out of 59 horses. This shows that individually (elicited or defined) adjectives are so far the most reliable method to assess individual differences in horses. The second finding is that horse personality traits can be grouped into three personality dimensions labelled as neuroticism, extroversion and agreeableness. This is in line with other horse personality research (Worth-Estes, 1952; Morris et al., 2002; Sian-Lloyd et al., 2006, in press) and with personality studies on other species (Eysenck, 1952; Costa & McCrae, 1992; Gosling, 2001). The third, most important, finding is that observers' empathy, emotional intelligence and two personality dimensions did not have any impact on 20 out if 26 correlations of horse personality ratings. This suggests that the observers did not project their own personalities onto horses.

Furthermore, this thesis contributes to the field of animal personality research on the theoretical as well as the methodological level. On the theoretical level, we summarised the parameters for a comprehensive personality theory in Chapter 2. These are important for the development of the theory of animal personality as well as for the development of a reliable and valid instrument of assessment. In Chapter 3 we further reviewed the most relevant human personality theories to be able to place animal personality studies in their theoretical context in Chapter 4. Animal, and more specifically horse, personality theories were also grouped according to their methodologies, experimental design and relevance to our project. We were intrigued by the question whether non-human animals have the capacity to assess others' personality (King et al., 1999; Morris et al., 2002). So we introduced a concept of animal-as-scientist based on Kelly's Personal Construct Psychology (1955) in Chapter 5. This theoretical model of animal behaviour integrates both quantitative and qualitative aspects of behaviour, including animal personality. The model does not inherit the mind-body problem of dualism, and has the capacity to incorporate animal consciousness, emotions and subjectivity by transposing them into animals' expressive behaviour (Wemelsfelder, 1997a, 1997b, 1999a, 2001; Dutton & Williams, 2004). We argued that like us, animals also strive to make sense of the world around them by construing events and making predictions or hypotheses. These events include the behaviour and personality of others. Hypotheses are tested in interactions. Based on these animals either modify their constructs or keep them unchanged. Our argument is that successful social animal-animal or human-animal interactions manifest animals' ability to assess others' consistent behaviours (Sanders, 1992). A method consistent with this idea is Kelly's Repertory grid technique, which we described in detail in Chapter 6. We included modifications we have made to adjust the method for horse personality assessment. In Chapter 7 we describe our first study, involving two stables and four different groups of observers. In the following study, reported in Chapter 8, we modified the experimental design by increasing the number of horses and familiar observers. We further unified the scoring lists to allow more comparisons between the observer groups. Both studies provided a pool of horse personality descriptors the observers used consistently. Finally, in Chapter 9 we report the effects the observers' individual differences had on their assessment of horses.

10.8. Project Limitations

Firstly, the personality ratings provided by familiar and unfamiliar raters were not significantly correlated, which could represent low validity of an assessment using repertory grid technique. However, it could also mean that by using video clips of horses in a human interaction test as opposed to personality assessment based on previous experience, we were measuring two different things. It is most likely that familiar observers were assessing personality or their relationships, while unfamiliar observers rated horses on the style of their interactions with others (Crossley, 1996). If this assumption is correct than the validity of the method has not been impaired. An additional argument in favour of this assumption is that our personality dimensions are in line with previous research on animal personality. However, when adopting trait classification of descriptors, we must be aware that one of the disadvantages of the trait model is the blurred distinction between more transient states and more stable personality traits (Plutchik, 1994; Matthews & Deary, 1998).

Secondly, personality traits are more likely to be expressed in a familiar environment (Matthews & Deary, 1998). Therefore familiar observers will see animals interacting in a familiar environment over a longer period of time. In our study videos were recorded in a fairly new environment and the horses were interacting with a complete stranger.

Thirdly, we were limited by our choice of horses and familiar observers. We had requirements for the number of horses at a stable and the number of participants who knew the horses well. Not many were available at the time our experiments took place.

Finally, all our familiar observers were women; therefore we decided that all the unfamiliar observers should also be females to minimize variation due to gender.

10.9. Application of the results and suggestions for further research

Application of our horse personality research is both theoretical and practical. On a theoretical level the results can help us to understand the biological and genetic basis of personality and its evolutionary value. We have shown that at least two personality dimensions that are consistent in different species, neuroticism and extroversion (Golsing, 2001), are also present in horses. On a practical level horses' personality is important in choosing individual animals for a specific purpose or environment (French, 1993). Humans use horses in the military, the police and sports, and in horse assisted therapy. Therefore horses' individual differences may predict how well a horse will fit each environment or purpose. This has implications for both horse performance and welfare. In our studies we have shown that there is a relationship between performance and personality and that more agreeable horses are also rated as better performers. In this respect, the relationship between the personality profiles of horses, their performance and their welfare should be further investigated. One of the applications of further investigation could be the development of an instrument that would help to place the right animal in the right environment, which would improve their performance and their welfare.

Methodologically, we have demonstrated that repertory grid technique is a reliable method for assessing horse personality. The pool of constructs we have generated in these two studies could be used as a platform for constructing a horse personality questionnaire, the measurement instrument mentioned above. Its reliability and validity should be tested further by cross validation with one of the well established personality questionnaires. To further address validity, the videos of horses in a human interaction test could be assessed by familiar observers.

Finally, we have introduced the theoretical concept of animal-as-scientist, which has the capacity to be developed as an animal behaviour model, incorporating quantitative and qualitative aspects of behaviour, including animal personality. The concept, however, needs be developed further theoretically and experimentally.

10.10. Conclusions

In conclusion, this project has discovered that repertory grid technique is a reliable method for horse personality assessment. It has identified three horse personality dimensions labelled as extroversion, neuroticism and agreeableness (Costa & McCrae, 1992). Finally, it has shown that observers do not project their own individual differences onto horses. However, these findings have uncovered other issues we still need to investigate: first we need to validate the repertory grid technique by cross validating the ratings with a well established personality questionnaire; secondly, we must develop a horse personality questionnaire based on the constructs generated in the two studies; thirdly we need to investigate the relationship between horse personality, performance and welfare; and finally, we need to further develop the idea of animal-as-scientist. In this project we have investigated horse personality. However, the method and the outcomes should be applied to other species as well to further strengthen both integrative methodologies and the research on animal individual differences.

References

Adams-Webber, J. R. (1970). Elicited versus provided constructs in repertory grid technique: a review. *Br. J. med. Psychol.*, *43*, 349-354.

Allen, C. & Bekoff, M. (1995). Cognitive ethology and the intentionality of animal behaviour. *Mind & Language, 10(4)*, 313-328.

Anderson, M. K., Friend, T. H., Evans, J. W., & Bushong, D. M. (1999). Behavioral assessment of horses in therapeutic riding programs. *Applied Animal Behaviour Science, 63*, 11-24.

Bekoff, M. (1998). Deep ethology, animal rights, and the great ape/animal project: resisting speciesism and expanding the community of equals. *Journal of Agricultural and Environmental Ethics, 10*, 269-296.

Bem, D. J. & Allen, A. (1974). On predicting some of the people some of the time: the search for cross-situational consistencies in behaviour. *Psychological Review, 81*, 506-520.

Birch, A. & Hayward, S. (1994). *Individual differences*. The Macmillan Press, London.

Boeree, C. G: George Kelly (2006). http://www.ship.edu/~cgboeree/kelly.html

Boyd, L. & Bandi, N. (2002). Reintroduction of takhi, *Equus ferus przewalskii*, to Hustai National Park, Mongolia: time budget and synchrony of activity pre- and post-release. *Applied Animal Behaviour Science, 78*, 87-102.

Boyd, L. & Houpt, K. A. (Eds.) (1994). *Przewalski's Horse. The History and Biology of An Endangered Species*. SUNY Press, Albany.

Bradshaw, W S (1992) *The Behaviour of the Domestic Cat*. CAB International: Wallingford.

Bradshaw, J. W. S. & Cook, S. E. (1996). Patterns of pet cat behaviour at feeding occasions. *Applied Animal Behaviour Science, 47*, 61-74.

Broom, D. M. (1991a). Assessing welfare and suffering. *Behavioural Processes, 25*, 117-123.

Broom, D. M. (1991b). Animal welfare: concepts and measurement. *Journal of animal science, 69,* 4167-4175.

Buckley P., Dunn T. & More S. J. (2004). Owners' perceptions of the health and performance of Pony Club horses in Australia. *Preventive Veterinary Medicine, 63*, 121-133

Budiansky, S. (1998). *The nature of horses: their evolution, intelligence and behaviour*. London : Phoenix Illustrated.

Buirski, P., & Plutchik, R. (1991). Measurement of deviant behaviour in a Gombe chimpanzee: Relation to later behavior. *Primates, 32*, 207-211.

Buss, D. M. (1989). Personality as traits. *American Psychologist 44,* 1378-1388.

Capitanio, J. P., Mendoza S. P. & Bentson K. L. (2004). Personality characteristics and basal cortisol concentrations in adult male rhesus macaques (*Macaca mulatta*). *Psychoneuroendocrinology, 29*, 1300-1308.

Cavalieri, P. & Singer, P. (eds) (1993). *The great ape project,* Fourth Estate: London.

Cattell, Raymond B. (1957). *Personality and motivation structure and measurement*. World Book Co.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2(3),* 200-220.

Chaplin, W. F., Oliver J. P. & Goldberg, L. R. (1988). Conceptions of States and Traits : Dimensional Attributes With Ideals as Prototypes. *Journal of Personality and Social Psychology, 54(4)*, 541-557.

Cheney, D. L. & Seyfarth, R. M. (1982). How monkeys perceive their grunts: Field playback experiments. *Animal Behaviour, 30*, 739-751.

Chomsky, N. (1972). *Gramatika i um*. Nolit: Beograd.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use.* Praeger Publishers: New York.

Clarke, A. S., & Boinski, S. (1995). Temperament in nonhuman primates. *American Journal of Primatology, 37*, 103-125.

Cooper, C 1998 *Individual differences*. Arnold: London.

Costa, P. T. & McCrae, R. R. (1987). Neuroticism, somatic complaints, and disease: Is the bark worse than the bite'? *Journal of Personality, 55,* 299-316.

Costa, P. T. & McCrae, R. R. (1992). *NEO Personality Inventory*. Psychological Assessment Resources, Inc.

Costa, P. T. & McCrae, R. R. (1995). Primary traits of Eysenck's P-E-N system: three and five factor solutions. *Journal of Personality and Social Psychology, 69*, 308-317.

Crist, Eileen (2000). *Images of Animals: Anthropomorphism and Animal Mind.* Philadelphia: Temple University Press.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: John Wiley.

Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.

Crossley, N. (1996). *Intersubjectivity: The Fabric of Social Becoming.* London: SAGE Publications.

Darwin, C. (1998). *The expression of emotions in man and animals.* Oxford: Oxford University Press. (Original work published, 1872).

Dawkins, M. S. (1980). *Animal Suffering*. London: Chapman and Hall.

Dawkins, M. S. (1988). Behaviour Deprivation: A Central Problem in Animal Welfare. *Applied Animal Behaviour Science, 20*, 209-225.

Dawkins, M. S. (1990). From an animal's point of view: Motivation, fitness, and animal welfare. *Behavioral and Brain Sciences, 13*, 1-61.

Dawkins, M. S. (1993). *Through Our Eyes Only?* Oxford: W. H. Freeman.

Deary, I. J. (1996). A (latent) big five personality model in 1915? A reanalysis of Webb's data. *Journal of Personality and Social Psychology, 71, 992-995.*

Descartes, R. (1989). Animals are machines. In T. Regan & P. Singer (Eds), *Animal Rights and Human Obligations* (pp. 13-19). New Jersey: Prentice-Hall.

Durr, R., & Smith, C. (1997). Individual differences and their relation to social structure in domestic cats. *Journal of Comparative Psychology, 111*, 412-418.

Dutton, D. (2006a, in press). "Destitute of Reason?" Mechanism, anthropomorphism and animal minds. In M. Bekoff (Ed.), *Encyclopedia of Human-Animal Relationships*. London & Westport, CT: Greenwood Press.

Dutton, D. (2006b, in press). The role of metaphors in constructing human-animal relationships. In M. Bekoff (Ed.), *Encyclopedia of Human-Animal Relationships*. London & Westport, CT: Greenwood Press.

Dutton, D.M. & Andersson, M. (2002). Personality in Royal Pythons and the human-snake relationship. *Anthrozoös, 15 (3)*, 243-250.

Dutton, D. M., Clark, R. A., & Dickins, D. W. (1997). Personality in captive chimpanzees: Use of a novel rating procedure. *International Journal of Primatology, 18,* 539-552.

Dutton, D. & Williams, C. (2004). A view from the bridge: subjectivity, embodiment and animal minds. *Anthrozoös, 17* (3), 210-224.

Worth-Estes, B. (1952). A study of the relationship between temperament of thoroughbred broodmares and performance of offspring. *The Journal of Genetic Psychology, 81*, 273-288.

Eysenck, H. J. (1952). *The scientific study of personality.* London: Routledge & Kegan Paul ltd.

Eysenck, H. J. (1997). Personality and experimental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology, 73,* 1224–1237.

Eysenck, M.W & Eysenck, H. J. (1980). Mischel and the concept of personality. *British Journal of Psychology, 71*, 191-204.

Eysenck 1991. Dimensions of personality: 16, 5 or 3?-Criteria for taxonomic paradigm. *Personality and Individual Differences, 12,* 773-790.

Feaver, J., Mendl, M., & Bateson, P. (1986). A method for rating the individual distinctiveness of domestic cats. *Animal Behaviour, 34*, 1016-1025.

Fjeld, S.P. & Landfield, A. W. (1961). Personal Construct Consistency. *Psychological Reports, 8,* 127-129.

Forkman, B., Furuhaug, I. L., & Jensen, P. (1995). Personality, coping patterns, and aggression in piglets. *Applied Animal Behaviour Science, 45,* 31-42.

Fouts, R. S. & Fouts, D. H. (1993). Chimpanzees' use of sign language. In P. Cavalieri & P. Singer (Eds.), *The great ape project* (pp. 28-41). London: Fourth Estate.

Fox, M. W. (1972). Socio-ecological implications of individual differences in wolf litters: A developmental and evolutionary perspective. *Behaviour, 41,* 298-313.

Fransella, F. & Bannister, D. (1977). *A manual for repertory grid technique.* London, New York: Academic Press.

Fransella, F, Bell, R. & Bannister, D. (2004). *A manual for repertory grid technique, second edition.* England: John Wiley & Sons.

Fraser, A. F. (1992). *The behaviour of the horse.* Wallingford: CAB International.

French, J. M. (1993). Assessment of donkey temperament and the influence of home environment. *Applied Animal Behaviour Science, 36*, 249-257.

Funder, D. C. (1995). On the accuracy of personality judgement: A realistic approach. *Psychological review, 102,* 652-670.

Funder, D. C. (1997). *The Personality Puzzle.* New York: Norton.

Funder, D. C. & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgement. *Journal of Personality and Social Psychology, 55,* 149-158.

Gains, N. (1994). The repertory grid approach. In H. J. H. MacFie & D. M. H. Thomson (Eds.), *Measurement of Food Preferences* (pp. 51-76). Blackie Academic and Professional.

Gains, N. & Thomson, D. M. H. (1990). Contextual evaluation of canned lagers using repertory grid method. *Int. Journal of Food Sci. Tech., 25,* 699-705.

Gaita, R. (2003). *The philosopher's dog.* London: Routledge

Gallup, G. G., Jr. (1970). Chimpanzees: Self-Recognition. *Science, 167*, 86-87.

Gallup, G. G., Jr. (1977). Self-recognition in primates: A comparative approach to the bidirectional properties of consciousness. *American Psychologist, 32,* 329-338.

Goddard, M. E., & Beilharz, R. G. (1983). Genetics of traits which determine the suitability of dogs as guide-dogs for the blind. *Applied Animal Ethology, 9*, 299-315.

Goddard, M. E., & Beilharz, R. G. (1984a). A factor analysis of fearfulness in potential guide dogs. *Applied Animal Behaviour Science, 12,* 253-265.

Goddard, M. E., & Beilharz, R. G. (1984b). The relationship of fearfulness to, and the effects of, sex, age and experience on exploration and activity in dogs. *Applied Animal Behaviour Science, 12*, 267-278.

Goddard, M. E., & Beilharz, R. G. (1986). Early prediction of adult behavior in potential guide dogs. *Applied Animal Behaviour Science, 15,* 247-260.

Gold, K. C., & Maple, T. L. (1994). Personality assessment in the gorilla and its utility as a management tool. *Zoo Biology, 13*, 509-522.

Goldberg, L. R. (1990). An alternative 'Description of Personality': the Big-Five factor structure. *Journal of Personality and Social Psychology, 59,* 1216-1229.

Goldberg (1993). *American Psychologist, 48*, 26.

Goodall, J. (1999). *In the Shadow of Man*. Phoenix Press.

Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin, 127,* 45-86.

Gosling, S. D., & Bonnenburg, A. V. (1998). Integrating personality research in anthrozoology: Ratings of six species of pets and their owners. *Anthrozoös, 11,* 148-156.

Gosling, S. D., & John, O. P. (1999). Personality dimensions in non-human animals: A cross-species review. Current Directions. *Psychological Science, 8,* 69-75.

Grajfoner, D., Wemelsfelder F. & Austin E. (2002). The qualitative assessment of pig behaviour using repertory grid technique, *BSAS Proceedings,* Annual Meeting in York.

Gray, J. A. (1978). The neuropsychology of anxiety. *British Journal of Psychology, 69,* 417-434.

Gray, J. A., Feldon, J., Rawlins, J., Helmsley, D. & Smith, A., 1991. The neuropsychology of schizophrenia. *Behavioural and Brain Sciences, 14*, 1–84.

Griffin, D. R. (1978). Prospects for a cognitive ethology. *The Behavioral and Brain Sciences, 4*, 527-629.

Griffin, D. R. (1985). Animal consciousness. *Neuroscience and Biobehavioural Reviews, 9*, 615-622.

Griffin, D. R. (1992). *Animal Minds*. Chicago: The University of Chicago Press.

Grignard, L., Boivin, X., Boissy, A. & Neindre, P. le (2001). Do beef cattle react consistently to different handling situations? *Applied Animal Behaviour Science, 71 (4)*, 263-276.

Hacker, P. M. S. (1997). *Wittgenstein: On human nature*. London: Phoenix.

Hafez, E. S. (Ed.) (1969). *The Behaviour of Domestic Animals.* Bailliere, Tindall & Cassell.

Hall, C. S. (1941). Temperament: a survey of animal studies. *Psychological Bulletin, 38*, 909-943.

Hampson, S. E. (1995). *The Construction of Personality.* London: Routledge.

Hanggi, E. B. (1999). Categorization learning in the horse (Equus caballus). *Journal of Comparative Psychology, 113(3)*, 1-10.

Hansen, S. W. (1996). Selection for behavioral traits in farm mink. *Applied Animal Behaviour Science, 49,* 137-148.

Hayne, S. M. & Gonyou, H. W. (2003). Effects of regrouping on the individual behavioural characteristics of pigs. *Applied Animal Behaviour Science, 82*, 267-278.

Hearne, V. (1994). *Adam's Task.* New York: HarperPerennial.

Hearne, V. (1995). *Animal Happiness.* New York: HarperPerennial.

Hebb, D. O. (1946). Emotions in man and animals: an analysis of the intuitive process of recognition. *Psychological Review. 53,* 88-106.

Hessing, M. J. C., Hagelsø, A. M., Van Beek, J. A., Wiepkema, P. R., Schouten, W. P. G., & Krukow, R. (1993). Individual behavioural characteristics in pigs. *Applied Animal Behaviour Science, 37*, 285-295.

Heyes, C. (1994). Reflections on self-recognition in primates. *Animal Behaviour, 47,* 909-919.

Hinde, R. A. & Spencer-Booth, Y. (1971). Toward understanding individual differences in rhesus motherinfant interaction. *Animal Behaviour, 19*, 165-173.

Hribar, T. (1993). *Fenomenologija I: Brentano, Heidegger, Husserl.* Ljubljana: Slovenska matica.

Husserl, E. (1970). *The crisis of European sciences and transcendental phenomenology.* Evanston, III: Northwestern University Press.

Hutson G. D. & Haskell M. J. (1997). Pre-race behaviour of horses as a predictor of race finishing order. *Applied Animal Behaviour Science, 53*, 231-248.

Jensen, P., Forkman, B., Thodberg, K. & Koster, E. (1995). Individual variation and consistency in piglet behaviour. *Applied Animal Behaviour Science, 45*, 43-52.

Kellerman, H. & Plutchik, R. (1968). Emotion-trait interrelations and the measurement of personality. *Psychological Reports, 23,* 1107-1114.

Kelly, G. A. (1955). *The psychology of personal constructs: A theory of personality, Volume One.* New York, London: W. W. Norton & Company, Inc.

Kelly, G. A. (1963). *A theory of personality: the psychology of personal constructs.* New York, London: W. W. Norton & Company.

Kenny, D. A. & Albright, L. (1987). Accuracy in Interpersonal Perception : A Social Relations Analysis. *Psychological Bulletin, 102(3)*, 390-402.

Kenny, D. A., Albright, L., Malloy, T. E. & Kashy, D. A. (1994). Consensus in Interpersonal Perception : Acquaintance and the Big Five. *Psychological Bulletin, 116(2),* 245-258.

Kerlinger, F. N. & Lee, H. B. (2000). *Foundations of Behavioural Research.* Orlando: Harcourt College Publishers.

Kiley-Worthington, M. (1987). *The behaviour of horses in relation to management and training.* London: J. A. Allen.

Kilgour, R. (1975). The open-field test as an assessment of the temperament of dairy cows. *Animal Behavior, 23,* 615-624.

King, J. E., Rumbaugh, D. M., & Savage-Rumbaugh, S. E. (1999). Perception of personality traits and semantic learning in evolving hominids. In M. C. Corballis, & S. E. G. Lea (Eds.), *The Descent of Mind* (pp. 80-97). Oxford: Oxford University Press.

Krechevsky, I. (1975). "Hypothesis" versus "chance" in the pre-solution period in sensory discrimination learning. In M. Levine (Ed.), *A cognitive theory of learning: research on hypothesis testing* (pp. 5-15). Hillsdale: Lawrence Erlbaum Associates.

Kruskal, J. B. & Wish, M. (1978). *Multidimensional scaling.* Beverly Hills: Sage.

Landfield, A. W. & Epting, F. R. (1987). *Personal construct psychology: clinical and personality assessment.* New York: Human Sciences Press, Inc.

Lansade, L., Bertrand, M., Boivin, X. & Bouissou, M. (2004). Effects of handling at weaning on manageability and reactivity of foals. *Applied Animal Science Behaviour Science, 87,* 131-149.

Lawrence, A. B., Terlouw, E. M. C., & Illius, A. W. (1991). Individual differences in behavioural responses of pigs exposed to non-social and social challenges. *Applied Animal Behaviour Science, 30*, 73-86.

LeDoux, J. (1998). *The Emotional Brain.* London: Phoenix.

Le Scolan, N., Hausberger, M., & Wolff, A. (1997). Stability over situations in temperamental traits of horses as revealed by experimental and scoring approaches. *Behavioural Processes, 41*, 257-266.

Levine, M. (1959). A model of hypothesis behaviour in discrimination learning set. *Psychological Review, 66*, 353-366.

Lilienfeld, S. O., Gershon, J., Duke, M., Marino, L., & de Waal, F. B. M. (1999). A preliminary investigation of the construct of psychopathic personality (psychopathy) in chimpanzees (Pan troglodytes). *Journal of Comparative Psychology, 113,* 365-375.

Lyons, D. M. (1989). Individual differences in temperament of dairy goats and the inhibition of milk ejection. *Applied Animal Behaviour Science, 22,* 269-282.

Macfie, H. J. H. & Thomson, D. M. H. (1994). *Measurement of food preferences.* London: Blackie Academic and Professional.

Mal, M. E., Friend, T. H., Lay, D. C. & Vogelsang, S. G. (1991). Behavioral responses of mares to short-term confinement and social isolation. *Applied Animal Behaviour Science, 31*, 13-24.

Mal, M. E., & McCall, C. A. (1996). The influence of handling during different ages on a halter training test in foals. *Applied Animal Behaviour Science, 50,* 115–120.

Martau, P. A., Caine, N. G., & Candland, D. K. (1985). Reliability of the Emotions Profile Index, primate form, with Papio hamadryas, Macaca fuscata, and two Saimiri species. *Primates, 26*, 501-505.

Martin, P. & Bateson, P. (1998). *Measuring behavior: An introductory guide (2nd ed.).* Cambridge: Cambridge University Press.

Mather, J. A., & Anderson, R. C. (1993). Personalities of Octopuses (Octopus rubescens). *Journal of Comparative Psychology, 107*, 336-340.

Matthews, G. & Deary, I. J. (1998). *Personality Traits.* Cambridge: Cambridge University Press.

Matthews, G. & K. Gilliland (1999). "The personality theories of H. J. Eysenck and J. A. Gray: a comparative review." *Personality and Individual Differences, 26*, 583–626.

Matthews, G., Jones, D.M. & Chamberlain, A.G. (1990). Refining the measurement of mood: the UWIST Mood Adjective Checklist. *British Journal of Psychology, 81,* 17-42.

McAfee, L. M. Mills, D. S. & Cooper, J. J. (2002). The use of mirrors for the control of stereotypic weaving behaviour in the stabled horse. *Applied Animal Behaviour Science, 78,* 159-173.

McCall, C.A., Hall, S., McElhenney, W.H. & Cummins, K.A. (2006). Evaluation and comparison of four methods of ranking horses based on reactivity. *Applied Animal Behaviour Science, 96*, 115-127.

McCall, C. A., Salters, M. A., Johnson, K. B., Silverman, S. J., McElhenney, W. H. & Lishak, R. S. (2003). Equine utilization of a previously learned visual stimulus to solve a novel task. *Applied Animal Behaviour Science, 82*, 163-172.

McCann, J. S., Heird, J. C., Bell, R. W., & Lutherer, L. O. (1988). Normal and more highly reactive horses: Heart rate, respiration rate and behavioral observations. *Applied Animal Behaviour Science, 19*, 201-214.

McCrae, R. R. & Costa, P. T. (1987). Validation of the Five-Factor Model of Personality Across Instruments and Observers. *Journal of Personality and Social Psychology, 52,* 81-90.

McCrae, R. R. & Costa, P. T. (1996). Personality trait structure as a human universal. *American Psychologist, 52*, 509-516.

McCrae, R. R. & Costa, P. T. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of personality and social psychology 78,* 173-186.

McCrae, R. R. & Costa, P. T. (2004). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences, 36*, 587-596

McFarland, D. (1990) *Animal behaviour*. Oxford: Oxford University Press.

Meddin, J. (1979). Chimpanzees, symbols, and the reflective self. *Social Psychology Quarterly, 42*, 99-109.

Mehrabian, A. & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40,* 525-543.

Merleau-Ponty, M. (1963). *The Structure of Behaviour.* Boston: Beacon Press.

Miles, H. L. (1997). Anthropomorphism, apes, and language. In R. W.Mitchell, N. S. Thompson, & H. L. Miles (Eds.), *Anthropomorphism, anecdotes and animals* (pp. 383-404). Albany: SUNY Press.

Mills, D. S. (1998). Personality and Individual differences in the horse, their significance, use and measurement. *Equine Veterinary Journal Supplement, 27*, 10-13.

Mischel, W. (1976). *Introduction to Personality*. Holt, Rinehart and Winston.

Momozawa, Y., Kusunose, R., Kikusui, T., Takeuchi Y. & Mori Y. (2005). Assessment of equine temperament questionnaire by comparing factor structure between two separate surveys, *Applied Animal Behaviour Science, 92*, pp. 77–84.

Momozawa, Y., Ono, T., Sato, F., Kikusui, T., Takeuchi, Y., Mori Y., & Kusunose, R., (2003). Assessment of equine temperament by a questionnaire survey to caretakers and evaluation of its reliability by simultaneous behaviour test, *Applied Animal Behaviour Science, 84*, 127–138.

Morgan, K., Zetterqvist, M., Hassmen, P., Visser, E. K., Rundgren, M, Blokhuis, H. J. (2000). Rider's personality and the perception of the co-operation between rider and horse. In *51th Annual Meeting of the European Association for Animal Production (EAAP)*. The Hague.

Morris, P.H., Gale A. & Duffy K. (2002). Can judges agree on the personality of horses? *Personality and Individual Differences, 33*, 67–81.

Murphy, J. A. (1995). Assessment of the temperament of potential guide dogs. *Anthrozoos, 8,* 224-228.

Murphy, J. A. (1998). Describing categories of temperament in potential guide dogs for the blind. *Applied Animal Behaviour Science, 58,* 163-178.

Murphy, J., Waldmann, T. & Arkins, S. (2004). Sex differences in equine learning skills and visuo-spatial ability. *Applied Animal Behaviour Science, 87,* 119-130.

Nagel, T. (1974). What is like to be a bat? *Philosophical Review,* 83, 435-451.

Nagel, T. (1986). *The View form Nowhere*. Oxford: Oxford University Press.

Overall, K. (1994). Temperament testing and training: Do they prevent behavioral problems? *Canine Practice, 19*, 19-21.

Overall, K. (1997). *Clinical Behavioural Medicine For Small Animals*. Mosby.

Panksepp, J. and Panksepp, J. B. (2000). The seven sins of evolutionary psychology. *Evolution and Cognition, 6*, 108-131.

Papaeliou, C. and Trevarthen, C. (1994): The infancy of music in Schuster, L. and Cambouropoulos E. (Eds.): *Musical Praxis Vol. 1(2)*: 19-33.

Paul, E. S. (2000). Empathy with animals and with humans: are they linked? *Anthrozoos, 13,* 194-202.

Pepperberg, I. M. (1983). Cognition in the African Grey parrot: Preliminary evidence for auditory/vocal comprehension of the class concept. *Animal Learning & Behaviour, 11,* 179-185.

Pervin, L. A. & John, O. P. (1997). *Personality: Theory and Research*. New York: John Wiley & Sons.

Piggins, D. & Phillips, C. J. C. (1998). Awareness in domesticated animals - concepts and definitions. *Applied Animal Behaviour Science, 57,* 181-200.

Plutchik, R. (1994). *The psychology and biology of emotion.* New York: HarperCollins College Publishers.

Plutchik, R., & Kellerman, H. (1974). *Emotions Profile Index.* Los Angeles, CA: Western Psychological Services.

Plutchik, R. & Kellerman, H. (Eds) (1980). *Emotion: Theory, Research, and Experience, Volume 1: Theories of Emotion.* New York: Academic Press.

Podberscek, A. L., & Gosling, S. D. (2000). Personality research on pets and their owners: Conceptual issues and review. In A. L. Podberscek, E. Paul, & J. A. Serpell (Eds.), *Companion animals and us: Exploring the relationships between people and pets* (pp. 143-167). Cambridge: Cambridge University Press.

Potter, J. T., Evans, J. W. & Nolt, B. H. (1994). Therapeutic horseback riding. *Journal of American Veterinary Medicine Association, 204*, 131-133.

Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioural and Brain Science, 1*, 515-526.

Priest, S. (1998). *Merleau-Ponty*. London: Routledge.

Rogers, C. (1989). *The Carl Rogers Reader*. Mariner Books.

Rosenthal, R & Rosnow, R. L. (1991). *Essentials of behavioural research: Methods and data analysis.* McGraw-Hill, Inc.

Ruefenacht, S., Gebhardt-Henrich, S., Miyake, T. & Gaillard, C. (2002). A behaviour test on German Shepherd dogs: Heritability of seven different traits. *Applied Animal Behaviour Science, 79*, 113–132.

Ruis, M. A. W., Brake, J. H. A., van de Burgwal, J. A., de Jong, I. C., Blokhuis, H. J., & Koolhaas, J. M. (2000). Personalities in female domesticated pigs: Behavioural and physiological indications. *Applied Animal Behaviour Science, 66*, 31-47.

Rychlak, J. F. (1981). *Personality and psychotherapy.* Boston: Houghton Mifflin Company.

Ryle, A. & Lunghi, M. (1970). The Dyad grid: A modification of repertory grid technique. *British Journal of Psychiatry, 117,* 323-327.

Sanders, C. R. (1992). Perception of intersubjectivity and the process of "speaking-for" in canine-human relationship. *Manuscript presented at the International Conference on "Science and the Human-Animal Relationship".* Amsterdam: The Netherlands.

Sanders, C. R. (1993). Understanding dogs: Caretakers' attributions of mindedness in canine-human relationships. *Journal of Contemporary Ethnography, 22*, 205-226.

Sanders, C. R. & Arluke, A. (1993). If Lions Could Speak. *The Sociological Quarterly, 34*, 377-390.

Saslow, C. A. (2002). Understanding the perceptual world of horses. *Applied Animal Behaviour Science, 78,* 209–224.

Schrader, L. (2002). Consistency of individual behavioural characteristics of dairy cows in their home pen. *Applied Animal Behaviour Science,* 77, 255–266.

Schutte, N.S., Malouff, J.M., Hall, L.E., Haggerty, D.J., Cooper, J.T., Golden, C.J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25*, 167-177.

Schutz, A. (1967). *The phenomenology of the social world.* Evanston: Northwestern University Press.

Scott, J. P. (1987). Why does human twin research not produce results consistent with those from nonhuman animals? *Behavioural & Brain Sciences, 10,* 39-40.

Seaman, S. C. Davidson, H. P. B. & Waran, N. K. (2002). How reliable is temperament assessment in the domestic horse (*Equus caballus*)? *Applied Animal Behaviour Science, 78,* 175–191.

Shapiro, K. J. (1989). The Death of the animal: Ontological vulnerability. *Between the Species, Fall 1989,* 183-194.

Shapiro, K. J. (1990). Understanding dogs through kinesthetic empathy, social construction, and history. *Anthrozoos, 3,* 184-195.

Sian Lloyd, A., Martin J. E., Bornett-Gauci, H. L. I. & Wilkinson R. G. (2006, in press). Evaluation of a novel method of horse personality assessment: Rater-agreement and links to behaviour. *Applied animal behaviour science, in press*.

Sinn, D. L. & Moltschaniwskyj, N. A. (2005). Personality traits in dumpling squid (*Euprymna tasmanica*): Context-specific traits and their correlation with biological characteristics. *Journal of Comparative Psychology,119*, 99-110.

Slater, P. J. B. (1981). Individual differences in animal behavior. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in ethology* (Vol. 4, pp. 35-49). New York: Plenum Press.

Smail, D. J. (1972). A grid measure of empathy in a therapeutic group. *British Journal of Medical Psychology, 45*, 165-169.

Sondergaard, E. & Halekoh, U. (2003). Young horses'reaction to humans in relation to handling and social environment. *Applied Animal Science Behaviour Science, 84*, 265-280.

Sondergaard, E. & Ladewig, J. (2004). Group housing exerts a positive effect on the behaviour of young horses during training. *Applied Animal Behaviour Science, 87,* 105-118.

Spencer-Booth, Y. & Hinde, R. A. (1969). Tests of behavioural characteristics for rhesus monkeys. *Behaviour, 33*, 180-211.

Spinath, F., Riemann, R., Hempel. S., Schlangen, B., Weiss, R., Borkenau, P. & Angleitner, A. (1999). A day in the life: description of the German observational study on adult twins (GOSAT) assessing twin similarity in controlled laboratory settings. In I. Mervielde, I. J. Deary, F. De Fruyt & F. Ostendorf (Eds.), *Personality Psychology in Europe* (pp. 311-328). Tilburg University Press.

Spoolder, H. A. M., Burbidge, J. A., Lawrence, A. B., Simmins, P. H., & Edwards, S. A. (1996). Individual behavioural differences in pigs: Intra- and inter-test consistency. *Applied Animal Behaviour Science, 49*, 185-198.

Stevenson-Hinde, J., & Zunz, M. (1978). Subjective assessment of individual rhesus monkeys. *Primates, 19,* 473-482.

Svartberg, K. (2002). Shyness-boldness predicts performance in working dogs. *Applied Animal Behaviour Science, 79*, 157-174.

Svartberg, K. & Forkman, B. (2002). Personality traits in the domestic dog (*Canis familiaris*). *Applied Animal Behaviour Science, 79*, 133-155.

Taylor Parker, S. & Gibson, K. R. (Eds) (1994). *"Language" and inteligence in monkeys and apes: Comparative developmental perspectives.* Cambridge: University of Cambridge.

Terrace, H. (1985). In the beginning was the "name". *American Psychologist, 40(9),* 1011-1028.

Tolman, E. C. & Krechevsky, I. (1933). Means-end-readiness and hypothesis – a contribution to comparative psychology. *Psychological Review, 40*, 60-70.

Trevarthen, C. (1979). The task of consciousness: how could the brain do them? *Brain and Mind, 69,* 187-252.

Trevarthen, C. (1985). Facial expressions of emotions in mother-infant interaction. *Human Neurobiology, 4.* 21-32.

Trevarthen, C. (1986): Development of intersubjective motor control in infants. In M. G. Wade and H. T. A. Whiting (eds.): *Motor Development in Children: Aspects of Coordination and Control*. Dordrecht, Martinus Nijhof, 209-261.

Trevarthen, C. (1990). Signs before speech. In T. A. Sebeok & J. Umiker-Sebeok (Eds.), *The semiotic web* (pp. 689-755

Trevarthen, C. (1993a). The function of emotions in early infant communication and development. In J. Nadel & L. Camaioni (Eds.), *New perspectives in early communicative development* (pp. 48-81). London: Routledge.

Trevarthen, C. (1993b). The self born in intersubjectivity: the psychology of an infant communicating. In U. Neisser (Ed.), *The Perceived Self: Ecological and Interpersonal Sources of Self-knowledge,* (pp. 121-173). Cambridge: Cambridge University Press.

Trevarthen, C. (2001). Intrinsic motives for companionship in understanding: their origin, development, and significance for infant mental health. *Infant Mental Health Journal, 22 (1-2)*, 95-131.

Trevarthen, C. & Aitken, K. J. (1994). Brain development, infant communication, and empathy disorders: Intrinsic factors in child mental health. *Development and Psychopathology, 6,* 597-633.

Tulloh, N. M. (1961). Behaviour of cattle: A study of temperament. *Animal Behaviour, 9,* 25-30.

Turner, D. C., Feaver, J., Mendl, M., & Bateson, P. (1986). Variation in domestic cat behaviour towards humans: a paternal effect. *Animal Behaviour, 34*, 1890-1892.

Visser, E. K., Rundgren, M., Zetterqvist, M., Morgan, K., Blokhuis, H. J. (2000). Do subjective ratings of riders correlate with the objective scores from behevioural tests? *51th Annual Meeting of the European Association for Animal Production (EAAP)*. The Hague.

Visser, E. K., van Reenen, C. G., Hopster, H., Schilder, M. B. H., Knaap, J. H., Barneveld, A. & Blokhuis, H.J. (2001). Quantifying aspects of young horses' temperament: consistency of behavioural traits. *Applied Animal Behaviour Science, 74,* 241–258.

Visser, E. K., van Reenen, C. G., van der Werf, J.T.N., Schilder, M. B. H., Knaap, J. H., Barneveld A., & Blokhuis, H. J. (2002). Heart rate and heart rate variability during a novel object test and a handling test in young horses, *Physiol. Behav., 76,* 289–296.

Visser, E. K., van Reenen, C. G., Schilder, M. B. H., Barneveld, A. & Blokhuis, H. J. (2003a). Learning preferences in young horse using two different learning tests. *Applied Animal Science Behaviour Science, 80,* 311-326.

Visser, E. K., van Reenen, C. G., Engel, B., Schilder, M. B. H., Barneveld, A. & Blokhuis, H. J. (2003b). The association between performance in show-jumping and personality traits earlier in life. *Applied Animal Behaviour Science, 82,* 279–295.

Waran N. K. & Cuddeford D. (1995). Effects of loading and transport on the heart rate and behaviour of horses. *Applied Animal Behaviour Science, 43*, 71-81.

Webster, J. (1994). *Animal welfare: A cool eye towards eden*. Oxford: Blackwell Science.

Weiss, A., King, J. E. & Perkins, L. (2006). Personality and subjective well-being in orangutans (*Pongo pygmaeus* and *Pongo abelii*). *Journal of Personality and Social Psychology, 90,* 501-511.

Wemelsfelder, F. (1985a). The question of animal awareness. In M. W. Fox et al. (Eds), *Advances in Animal Welfare Science* (pp. 1-18). Dordrecht: Martinus Nijhoff Publishers.

Wemelsfelder, F. (1985b). Animal boredom: Is scientific study of the subjective experiences of animals possible? In M. W. Fox et al. (Eds), *Advances in Animal Welfare Science* (pp. 115-154). Dordrecht: Martinus Nijhoff Publishers.

Wemelsfelder, F. (1993). *Animal Boredom: towards an empirical approach of animal subjectivity*. PhD Thesis, University of Leiden.

Wemelsfelder, F. (1997a). The scientific validity of subjective concepts in models of animal welfare. *Applied Animal Behaviour Science, 53*, 75-88.

Wemelsfelder, F. (1997b). Investigating the animal's point of view. An enquiry into a subject based method of measurement in the field of animal welfare. In M. Dol et al. (Eds.), *Animal Consciousness and Animal Ethics* (pp. 73-89). Assen: Van Gorcum.

Wemelsfelder, F. (1999a). The problem of animal subjectivity and its consequences for the scientific measurement of animal suffering. In F. L. Dolins (Ed.), *Attitudes to Animals: Views in Animal Welfare* (pp. 37-53). Cambridge: Cambridge University Press.

Wemelsfelder, F. (1999b). Anthropomorphism, anecdotes, and animals (Book review). *Animal Welfare, 8,* 185-187.

Wemelsfelder, F. (2001). The inside and outside aspects of consciousness: complementary approaches to the study of animal emotion. *Animal Welfare, 10*, S129-139.

Wemelsfelder, F., Hunter, E. A., Mendl, M. T., & Lawrence, A. B. (2000). The spontaneous qualitative assessment of behavioral expressions in pigs: First explorations of a novel methodology for integrative animal welfare measurement. *Applied Animal Behaviur Science, 67*, 193-215.

Wilson, D. S., Coleman, K., Clark, A. B., & Biederman, L. (1993). Shy-bold continuum in pumpkinseed sunfish (Lepomis gibbosus): An ecological study of a psychological trait. *Journal of Comparative Psychology, 107*, 250-260.

Winskill L. C., Waran N. K. & Young R. J.(1996). The effect of a foraging device (a modified 'Edinburgh Foodball') on the behaviour of the stabled horse. *Applied Animal Behaviour Science, 48,* 25-35.

Witkin, H. A. (1940). "Hypotheses" in rats: an experimental critique. *The Journal of Comparative Psychology, 30,* 457-482.

Wolff, A., Hausberger, M., & Le Scolan, N. (1997). Experimental tests to assess emotionality in horses. *Behavioural Processes, 40*, 209-221.

Wood-Gush, D. G. M. & Galbraith (David Grainger Marcus). (1987). Social relationships in a herd of 11 geldings and two female ponies. *British Veterinary Association*.

Wood-Gush, D. G. M. *et al.* (Eds.) (1980). *Self Awareness in Domesticated Animals*. Potters Bar: UFAW.

Zuckerman, M. (1991). *Psychobiology of Personality.* Cambridge: Cambridge University Press.

Zuckerman, M. (1992). What is a basic factor and which factors are basic? Turtles and the way down. *Personality and Individual Differences, 13,* 675-681.

APPENDICES

Appendix 1

Elicitation form with elaborations – Study 1, Stable 1 – familiar observers

OBSERVER 1			
EXPLANATION	**EMERGING**	**OPPOSITE**	**EXPLANATION**
Not stressed by life	**LAID BACK**	**STRESSFUL**	**Find life difficult**
Willing to cooperate	WILLING TO WORK	UNWILLING TO WORK	Not keen to apply oneself
Would be full of trust	**TRUSTFUL**	**NOT TRUSTWORTHY**	**Unreliable**
Grown up/ learned many things	**MATURE**	**IMMATURE**	**Still to learn/ to grow up**
Soft touch + approach	GENTLE	LESS GENTLE	Not much empathy
	RELIABLE	UNRELIABLE	Not yet to be trusted
Finds life easy	**COPES WELL**	**GET STRESSFUL**	**Finds work and life difficult**
Likes own company	LONER	LIKES COMPANY	Likes to be together
Shows authority	BULLY	SHY	
Likes to be in charge	**BOSSY**	**LIKES TO FOLLOW/ NOT AGGRESSIVE**	
	WORRIER	LAID BACK	
Can understand + learn	INTELLIGENT	UNABLE TO LEARN	Has problems learning
	NERVOUS	CONFIDENT	
Very reliable/ does a good job	**CONSCIENTIOUS**	**UNWILLING TO LEARN**	
Very sensible to riders needs	**GENUINE**	**FALSE**	**Unwilling to cooperate or learn**
Aware of needs	**SENSITIVE**	**TOUCH/ ABLE TO COPE**	
Not confident	UNSURE	CONFIDENT	Knows the job he does well
Very sure	RELIABLE	NOT VERY TRUSTING	Suspicious of people
Very genuine	NOT TROUBLE MAKERS	OUT TO MAKE TROUBLE/ MAKE A STATEMENT	Draw attention to themselves
	EASY TO LEAD	UNHELPFUL WHEN BEING LED	
	GOOD LEAD HORSE	NOT A NATURE BORN LEADER	Likes to follow
	QUICK TO LEARN	SLOW UNDERSTANDING	Needs time to learn
	WISE	**UNWISE DUE TO LACK OF EXPERIENCE**	
	CONFIDENT	NERVOUS	
To have empathy	**PERCEPTIVE**	**NOT ABLE TO RELATE**	**Unable to communicate**
	UPTIGHT	RELAXED	
	RELAXED	TENSE	
	IMPATIENT	**PATIENT/ WILLING TO LEARN**	**Willing to listen + learn**
	LIVELY	LAZY	Not willing to partake in work
	UNSURE	CONFIDENT	
	FRIENDLY	NERVOUS	
	STRESSED	RELAXED	
Does not trust	**DISTRUSTFUL**	**TRUSTING OF OTHERS**	**Accepting of others + themselves, trusting**

Does not like to join in	**WITHDRAWN**	**VERY OPEN**	**Loves to be involved with others**
Does not respond to surroundings	**UNRESPONSIVE**	**RESPONSIVE TO OUTSIDE FORCES**	**Outside + external stimulation**
Not to have stimulation	BOREDOM	ACTIVE, HAPPY, STIMULATED	
Full of joy	**HAPPY**	**SAD, DEPRESSED**	**Needs to find focus**

Appendix 2

Scoring form for Stable 1 – Familiar observers

CONFIDENT	1	2	3	4	5	UNSURE
HAPPY	1	2	3	4	5	**UNHAPPY**
LAID-BACK	1	2	3	4	5	**WORRIER**
RELIABLE	1	2	3	4	5	**UNRELIABLE**
BULLY	1	2	3	4	5	**BULLIED**
INTELLIGENT	1	2	3	4	5	**STUPID**
WILLING TO WORK	1	2	3	4	5	**UNWILLING TO WORK**
STEADY	1	2	3	4	5	**FAST**
BOLD	1	2	3	4	5	**TIMID**
FRIENDLY	1	2	3	4	5	**UNFRIENDLY**
FORWARD GOING	1	2	3	4	5	**SLOW**
DOES NOT LIKE SMALL AREAS	1	2	3	4	5	**NOT WORRIED ABOUT IT**
PUSHY	1	2	3	4	5	**DOCILE**
TRUSTFUL	1	2	3	4	5	**SUSPICIOUS**
AFFECTIONATE	1	2	3	4	5	**ALOOF**
BOSSY	1	2	3	4	5	**LIKES TO FOLLOW**
OBEDIENT	1	2	3	4	5	**DISOBEDIENT**
EASY TO LEAD	1	2	3	4	5	**STUBBORN**
GOOD LEAD HORSE	1	2	3	4	5	**PREFERS FOLLOWING**
INQUISITIVE	1	2	3	4	5	**NOT INQUISITIVE**
RESPONSIVE	1	2	3	4	5	**UNRESPONSIVE**
ACCEPTING	1	2	3	4	5	**NOT ACCEPTING**
SOCIABLE	1	2	3	4	5	**LONER**
LIVELY	1	2	3	4	5	**SLUGGISH**
PLAYFUL	1	2	3	4	5	**DULL**
QUIET	1	2	3	4	5	**EXCITABLE**
PATIENT	1	2	3	4	5	**IMPATIENT**
GOOD NATURED	1	2	3	4	5	**STROPPY**
RELAXED	1	2	3	4	5	**UPTIGHT**
ALERT	1	2	3	4	5	**APATHETIC**
SENSITIVE	1	2	3	4	5	**UNFEELING**
GREEDY	1	2	3	4	5	**NOT GREEDY**

Appendix 3

Scoring form for Stable 2 – Familiar observers

BOLD	1	2	3	4	5	**SHY**
CLEVER	1	2	3	4	5	**STUPID**
SOCIABLE	1	2	3	4	5	**ANTI-SOCIAL**
CHEEKY	1	2	3	4	5	**TIMID**
FRIENDLY	1	2	3	4	5	**UNFRIENDLY**
CALM	1	2	3	4	5	**EXCITABLE**
LAID-BACK	1	2	3	4	5	**HIGHLY STRUNG**
AFFECTIONATE	1	2	3	4	5	**ALOOF**
CONFIDENT	1	2	3	4	5	**NERVOUS**
WILLING TO PLEASE	1	2	3	4	5	**STUBBORN**
SAFE	1	2	3	4	5	**SPOOKY**
CONTENT	1	2	3	4	5	**UNSETTLED**
OBEDIENT	1	2	3	4	5	**DISOBEDIENT**
INTERESTED	1	2	3	4	5	**DISINTERESTED**
LIVELY	1	2	3	4	5	**DULL**
EASY TO WORK WITH	1	2	3	4	5	**DIFFICULT**
PLAYFUL	1	2	3	4	5	**BORING**
SECURE	1	2	3	4	5	**INSECURE**
PREDICTABLE	1	2	3	4	5	**UNPREDICTABLE**
HAPPY	1	2	3	4	5	**UNHAPPY**
WELL BEHAVED	1	2	3	4	5	**NAUGHTY**
SENSITIVE	1	2	3	4	5	**UNRESPONSIVE**
FRESH	1	2	3	4	5	**LAZY**
WELL MANNERED	1	2	3	4	5	**IGNORANT**
QUIET	1	2	3	4	5	**IMPATIENT**
GENTLE	1	2	3	4	5	**AGGRESSIVE**
RELIABLE	1	2	3	4	5	**UNTRUSTWORTHY**
GOOD NATURED	1	2	3	4	5	**STROPPY**

Appendix 4
Study 1 – Video observation, experienced observers,

OBSERVER 2			
EXPLANATION	**EMERGING**	**OPPOSITE**	**EXPLANATION**
	BOLD	**TIMID**	
Determined in his terms	BUSINESS-LIKE	APATHETIC	
Laid back	**UNPERTURBED**	**NERVOUS OF THE SITUATION**	
Unaware/ disinterested	SELF IMPORTANT	SHY	Aware
	VERY SURE OF HIMSELF	**TIMID**	
	LAID-BACK	ENERGETIC/BOLD	
	QUIET	BOLD	
	NOT VERY AWARE	AWARE	
	DISINTERESTED	**INTERESTED**	
	CHEEKY	POLITE	**Boring, unintelligent, shy**
	BRAVE	TIMID	
	RUDE	**POLITE**	
Demanding attention	DEMANDING	UNDEMANDING	
	PUSHY	POLITE	
Determined – doing what they feel	WILLFUL	BORING	
	ENERGETIC	QUIET	
	SHY	BOLD	
	UNCERTAIN	BRAVE	
	WORRIED	CONFIDENT	
	FRIENDLY	UNFRIENDLY/ COLD	
Seek people, social contact	CLINGY	INDEPENDENT/ INDIVIDUAL	
	NERVOUS	BOLD/ CONFIDENT	
Worried	**ANXIOUS**	**CONFIDENT**	
	AFRAID	BRAVE	
	UNSURE OF PEOPLE	CONFIDENT	
Suspicious	**UNTRUSTING**	**TRUSTING**	
	CAUTIOUS	BOLD	
Interested but not doing anything	**CURIOUS**	**DISINTERESTED**	
	AWARE / INTERESTED	DISINTERESTED	
	LAID-BACK	**ENERGETIC/ BRIGHT**	
	UNCONCERNED	WORRIED	
	WORRIED	CONFIDENT	
Inquisitive+doing	EXPLORATIVE	DISINTERESTED	
	CAUTIOUS	BOLD	
Intelligent/ sensitive	**SHARP**	**UNINTELLIGENT**	**Quiet, laid back**
Not doing bad	**KIND**	**UNKIND**	**Rude**
	DEPENDABLE	**NOT DEPENDABLE**	

Sensible	**WISE**	**NAÏVE/ UNINTELLIGENT**	
	BORED	**INTERESTED/ BRIGHT**	**Energetic**
	WARY	BOLD	
	WARY OF CONTACT	**BOLD/ APPROACHABLE**	
	BOSSY	**TIMID/ SUBDUED**	**Disinterested, quiet, given up**
Trustworthy	**SENSIBLE**	**SILLY**	**Unpredictable, spooky**
	FUN	BORING	
Happy with the way they are, not extreeme	**COMFORTABLE**	**WORRIED/ TENSE**	

Appendix 5

Scoring form for Experienced observers

INTERESTED	1	2	3	4	5	DISINTERESTED
BOLD	1	2	3	4	5	TIMID
FRIENDLY	1	2	3	4	5	UNFRIENDLY
CONFIDENT	1	2	3	4	5	NERVOUS
BORED	1	2	3	4	5	ALERT
RELAXED	1	2	3	4	5	TENSE
LAID-BACK	1	2	3	4	5	SPIRITED
CALM	1	2	3	4	5	EXCITED
UNCONCERNED	1	2	3	4	5	WORRIED
CONTENT	1	2	3	4	5	UNHAPPY
SOCIABLE	1	2	3	4	5	UNSOCIABLE
ACTIVE	1	2	3	4	5	PASSIVE
SECURE	1	2	3	4	5	FRIGHTENED
SPOOKY	1	2	3	4	5	TRUSTWORTHY
SURE	1	2	3	4	5	UNSURE
QUIET	1	2	3	4	5	ENERGETIC
TRUSTING	1	2	3	4	5	UNTRUSTING
PLACID	1	2	3	4	5	HOT TEMPERED
SAFE	1	2	3	4	5	JUMPY
DOCILE	1	2	3	4	5	HIGHLY STRUNG
SETTLED	1	2	3	4	5	UNSETTLED
PLAYFUL	1	2	3	4	5	SERIOUS
AWARE	1	2	3	4	5	OBLIVIOUS
GENTLE	1	2	3	4	5	AGGRESSIVE
AFFECTIONATE	1	2	3	4	5	UNAFFECTIONATE
STEADY	1	2	3	4	5	FLIGHTY
GOOD NATURED	1	2	3	4	5	DIFFICULT
INTELLIGENT	1	2	3	4	5	STUPID
EASY GOING	1	2	3	4	5	STUBBORN
EXPERIENCED	1	2	3	4	5	INEXPERIENCED
PATIENT	1	2	3	4	5	IMPATIENT

Appendix 6

Scoring form for Inexperienced observers

INTERESTED	1	2	3	4	5	**DISINTERESTED**
FRIENDLY	1	2	3	4	5	**UNFRIENDLY**
CALM	1	2	3	4	5	**AGITATED**
CONFIDENT	1	2	3	4	5	**NERVOUS**
HAPPY	1	2	3	4	5	**UNHAPPY**
PLACID	1	2	3	4	5	**HIGHLY STRUNG**
PLAYFUL	1	2	3	4	5	**SERIOUS**
ACTIVE	1	2	3	4	5	**PASSIVE**
SURE	1	2	3	4	5	**UNSURE**
BORED	1	2	3	4	5	**INTERESTED**
DOCILE	1	2	3	4	5	**SPIRITED**
ALERT	1	2	3	4	5	**TIRED**
INTERACTIVE	1	2	3	4	5	**DISTANT**
TOLERANT	1	2	3	4	5	**INTOLERANT**
RELAXED	1	2	3	4	5	**TENSE**
SECURE	1	2	3	4	5	**INSECURE**
TRUSTING	1	2	3	4	5	**DISTRUSTFUL**
EXTROVERTED	1	2	3	4	5	**INTROVERTED**
APPROACHABLE	1	2	3	4	5	**STAND-OFFISH**
AFFECTIONATE	1	2	3	4	5	**UNAFFECTIONATE**
BOLD	1	2	3	4	5	**TIMID**
ACCEPTING	1	2	3	4	5	**REJECTING**
SOCIABLE	1	2	3	4	5	**UNSOCIABLE**
PATIENT	1	2	3	4	5	**IMPATIENT**
GENTLE	1	2	3	4	5	**HARSH**
WELL BEHAVED	1	2	3	4	5	**STUBBORN**

Appendix 7

STABLE 1

Familiar raters scoring horses on all constructs

HORSE	W	χ^2	P
1. SAM	0.319 NS	39.504	0.141
2. HAMISH	0.592***	73.441	0.000
3. GINGER	0.674***	83.568	0.000
4. SANDIE	0.623***	77.300	0.000
5. CUILLIN	0.588***	72.960	0.000
6. STROLLER	0.446**	55.249	0.005
7. ZERO	0.420**	52.133	0.010
8. BOBBY	0.558***	69.136	0.000
9. EDDIE	0.636***	78.911	0.000
10. FUDGE	0.381*	47.298	0.031

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Experienced raters scoring horses on all constructs

HORSE	W	χ^2	P
1. SAM	0.328***	177.344	0.000
2. HAMISH	0.546***	294.881	0.000
3. GINGER	0.513***	277.146	0.000
4. SANDIE	0.387***	209.065	0.000
5. CUILLIN	0.607***	327.516	0.000
6. STROLLER	0.339***	182.964	0.000
7. ZERO	0.580***	312.992	0.000
8. BOBBY	0.514***	277.541	0.000
9. EDDIE	0.441***	237.943	0.000
10. FUDGE	0.495***	267.196	0.000

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Inexperienced raters scoring horses on all constructs

HORSE	W	χ^2	P
1. SAM	0.333***	141.377	0.000
2. HAMISH	0.351***	148.985	0.000
3. GINGER	0.549***	233.475	0.000
4. SANDIE	0.298***	126.707	0.000
5. CUILLIN	0.469***	199.254	0.000
6. STROLLER	0.311***	132.062	0.000
7. ZERO	0.511***	217.381	0.000
8. BOBBY	0.511***	217.079	0.000
9. EDDIE	0.229***	97.221	0.000
10. FUDGE	0.332***	140.987	0.000

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Appendix 8

STABLE 2

Familiar raters scoring horses on all constructs

HORSE	W	χ^2	P
1. LIZZY	0.695***	75.057	0.000
2. APPLEY	0.731***	78.968	0.000
3. MOLLY	0.586***	63.337	0.000
4. COCO	0.489**	52.806	0.002
5. LYRIC	0.440**	47.540	0.009
6. STROPPY	0.393*	42.453	0.030
7. REIVER	0.644***	69.516	0.000
8. SLOTTY	0.330 NS	35.615	0.124
9. MISTY	0.411*	44.408	0.019
10. ROB	0.535***	57.764	0.001
11. DOLCE	0.641***	69.265	0.000

Experienced raters scoring horses on all constructs

HORSE	W	χ^2	P
1. LIZZY	0.650***	370.432	0.000
2. APPLEY	0.494***	281.415	0.000
3. MOLLY	0.284***	162.114	0.000
4. COCO	0.313***	178.190	0.000
5. LYRIC	0.498***	284.040	0.000
6. STROPPY	0.165***	93.943	0.000
7. REIVER	0.323***	184.093	0.000
8. SLOTTY	0.240***	136.587	0.000
9. MISTY	0.409***	232.971	0.000
10. ROB	0.335***	191.107	0.000

Inexperienced raters scoring horses on all constructs

HORSE	W	χ^2	P
1. LIZZY	0.449***	190.753	0.000
2. APPLEY	0.605***	257.235	0.000
3. MOLLY	0.288***	122.475	0.000
4. COCO	0.310***	131.916	0.000
5. LYRIC	0.501***	213.047	0.000
6. STROPPY	0.305***	129.656	0.000
7. REIVER	0.252***	107.132	0.000
8. SLOTTY	0.315***	133.666	0.000
9. MISTY	0.372***	158.286	0.000
10. ROB	0.470***	199.580	0.000

Appendix 9

Difference between the constructs elicited by the majority (at least three observers) and partially provided constructs – Stable 1

HORSE	Familiar W		Experienced W		Inexperienced W	
	Majority	Provided	Majority	Provided	Majority	Provided
1. HAMISH	0.590**	0.573**	0.536***	0.578***	0.432***	0.254***
2. STROLLER	0.535*	0.399 NS	0.261***	0.445***	0.356***	0.298***
3. CUILLIN	0.646**	0.553**	0.610***	0.641***	0.569***	0.380***
4. SAM	0.466*	0.212 NS	0.441***	0.215***	0.284***	0.309***
5. BOBBY	0.555**	0.599***	0.533***	0.508***	0.632***	0.345***
6. ZERO	0.573**	0.298 NS	0.548***	0.609***	0.572***	0.414***
7. FUDGE	0.341 NS	0.393 NS	0.419***	0.588***	0.427***	0.226***
8. GINGER	0.690***	0.656***	0.507***	0.556***	0.694***	0.341***
9. SANDIE	0.682***	0.585***	0.432***	0.315***	0.351***	0.260***
10. EDDIE	0.644**	0.644***	0.397***	0.523***	0.318***	0.131**

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Difference between the constructs elicited by the majority (at least three observers) and partially provided constructs – Stable 2

HORSE	Familiar W		Experienced W		Inexperienced W	
	Majority	Provided	Majority	Provided	Majority	Provided
1. LIZZY	0.670**	0.720***	0.626***	0.654***	0.545***	0.321***
2. APPLEY	0.660**	0.734***	0.543***	0.398***	0.639***	0.563***
3. MOLLY	0.468*	0.511**	0.269***	0.305***	0.310***	0.278***
4. COCO	0.534*	0.467*	0.248***	0.398***	0.383***	0.234***
5. LYRIC	0.577**	0.350 NS	0.509***	0.464***	0.595***	0.398***
6. STROPPY	0.368 NS	0.418*	0.127**	0.195***	0.342***	0.289***
7. REIVER	0.730***	0.477*	0.380***	0.256***	0.368***	0.137**
8. SLOTTY	0.254 NS	0.263 NS	0.170***	0.274***	0.338***	0.307***
9. MISTY	0.439 NS	0.411 NS	0.510***	0.278***	0.464***	0.218***

10. ROB	0.522*	0.552**	0.306***	0.382***	0.562***	0.352***
12. DOLCE (absent when filming)	0.624**	0.688***				

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Appendix 10

Difference between the most useful constructs and the rest – Stable 1

HORSE	Familiar W		Experienced W		Inexperienced W	
	Most	Rest	Most	Rest	Most	Rest
1. SAM	0.383 NS	0.226 NS	0.414***	0.279***	0.339***	0.174**
2. HAMISH	0.503**	0.617**	0.569***	0.553***	0.263***	0.371***
3. GINGER	0.708***	0.597**	0.532***	0.509***	0.617***	0.418***
4. SANDIE	0.607***	0.656***	0.377***	0.381***	0.308***	0.231***
5. CUILLIN	0.547**	0.649**	0.622***	0.598***	0.503***	0.350***
6. STROLLER	0.485*	0.393 NS	0.341***	0.339***	0.217***	0.345***
7. ZERO	0.362 NS	0.385 NS	0.556***	0.602***	0.525***	0.315***
8. BOBBY	0.259 NS	0.634**	0.464***	0.563***	0.521***	0.534***
9. EDDIE	0587***	0.614**	0.400***	0.482***	0.217***	0.282***
10. FUDGE	0.364 NS	0.395 NS	0.460***	0.521***	0.386***	0.269***

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Difference between the most useful constructs and the rest – Stable 2

HORSE	**Familiar W**		**Experienced W**		**Inexperienced W**	
	Most	Rest	Most	Rest	Most	Rest
1. LIZZY	0.723***	0.628***	0.654***	0.658***	0.464***	0.356***
2. APPLEY	0.816***	0.569**	0.494***	0.455***	0.629***	0.500***
3. MOLLY	0.646**	0.621**	0.275***	0.279***	0.332***	0.328***
4. COCO	0.475*	0.506*	0.241***	0.388***	0.243***	0.339***
5. LYRIC	0.389 NS	0.515*	0.520***	0.468***	0.525***	0.449***
6. STROPPY	0.473*	0.354 NS	0.121**	0.196***	0.276***	0.345***
7. REIVER	0.566**	0.704***	0.368***	0.292***	0.283***	0.231***
8. SLOTTY	0.273 NS	0.388 NS	0.145***	0.298***	0.281***	0.358***
9. MISTY	0.291 NS	0.538**	0.506***	0.301***	0.399***	0.361***
10. ROB	0.458*	0.576**	0.305***	0.375***	0.489***	0.469***
11. DOLCE	0.628**	0.588**				

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Appendix 11

Consistency of familiar observers rating horses on separate constructs – Stable 1. Constructs with W .6 or above are in bold.

CONSTRUCT	W	χ^2	P
CONFIDENT-UNSURE	**0.697****	**25.105**	**0.003**
HAPPY-UNHAPPY	**0.637****	**22.943**	**0.006**
LAID-BACK-WORRIER	**0.815*****	**29.342**	**0.001**
RELIABLE-UNRELIABLE	**0.773*****	**27.817**	**0.001**
BULLY-BULLIED	0.198 NS	7.124	0.624
INTELLIGENT-STUPID	0.562*	20.225	0.017
WILLING TO WORK-UNWILLING TO WORK	0.502*	18.080	0.034
STEADY-FAST	0.591*	21.288	0.011
BOLD-TIMID	0.509*	18.309	0.032
FRIENDLY-UNFRIENDLY	0.521*	18.751	0.027
FORWARD GOING-SLOW	**0.674****	**24.277**	**0.004**
DOES NOT LIKE SMALL AREAS - NOT WORRIED ABOUT IT	0.375 NS	13.510	0.141
PUSHY-DOCILE	0.430 NS	15.488	0.078
TRUSTFUL-SUSPICIOUS	**0.608****	**21.906**	**0.009**
AFFECTIONATE-ALOOF	0.372 NS	13.407	0.145
BOSSY-LIKES TO FOLLOW	0.469 NS	16.871	0.051
OBEDIENT-DISOBEDIENT	**0.704****	**25.332**	**0.003**
EASY TO LEAD-STUBBORN	0.586*	21.110	0.012
GOOD LEAD HORSE – PREFERS FOLLOWING	**0.631****	**22.718**	**0.007**
INQUISITIVE-NOT INQUISITIVE	0.495*	17.828	0.037
RESPONSIVE-UNRESPONSIVE	0.473*	17.019	0.048
ACCEPTING-NOT ACCEPTING	0.576*	20.728	0.014
SOCIABLE-LONER	0.328 NS	11.802	0.225
LIVELY-SLUGGISH	**0.808*****	**29.087**	**0.001**
PLAYFUL-DULL	0.563*	20.259	0.016

QUIET-EXCITABLE	**0.733****	**26.400**	**0.002**
PATIENT-IMPATIENT	0.456 NS	16.431	0.058
GOOD NATURED-STROPPY	0.288 NS	10.367	0.322
RELAXED-UPTIGHT	**0.644****	**23.190**	**0.006**
ALERT-APATHETIC	0.453 NS	16.309	0.061
SENSITIVE-UNFEELING	0.267 NS	9.626	0.382
GREEDY-NOT GREEDY	**0.605****	**21.778**	**0.001**

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Appendix 12

Consistency of familiar observers rating horses on separate constructs – Stable 2

CONSTRUCT	W	χ^2	P
BOLD – SHY	0.592**	23.671	0.009
CLEVER – STUPID	**0.682****	**27.261**	**0.002**
SOCIABLE - ANTI-SOCIAL	**0.624****	**24.955**	**0.005**
CHEEKY – TIMID	**0.713*****	**28.537**	**0.001**
FRIENDLY – UNFRIENDLY	**0.664****	**26.579**	**0.003**
CALM – EXCITABLE	0.512*	20.460	0.025
LAID-BACK – HIGHLY STRUNG	**0.605****	**24.183**	**0.007**
AFFECTIONATE – ALOOF	**0.737*****	**29.476**	**0.001**
CONFIDENT – NERVOUS	**0.656****	**26.234**	**0.003**
WILLING TO PLEASE – STUBBORN	0.580**	23.204	0.010
SAFE – SPOOKY	**0.677****	**27.096**	**0.003**
CONTENT – UNSETTLED	0.542*	21.689	0.017
OBEDIENT – DISOBEDIENT	**0.659****	**26.346**	**0.003**
INTERESTED – DISINTERESTED	0.531*	21.259	0.019
LIVELY – DULL	**0.702****	**28.067**	**0.002**
EASY TO WORK WITH – DIFFICULT	0.390 NS	15.606	0.111
PLAYFUL – BORING	0.531*	21.250	0.019
SECURE – INSECURE	0.597**	23.872	0.008
PREDICTABLE – UNPREDICTABLE	0.474*	18.954	0.041
HAPPY – UNHAPPY	0.579**	23.167	0.010
WELL BEHAVED – NAUGHTY	**0.642****	**25.698**	**0.004**
SENSITIVE – UNRESPONSIVE	0.486*	19.458	0.035
FRESH – LAZY	0.451 NS	18.046	0.054
WELL MANNERED – IGNORANT	**0.704****	**28.157**	**0.002**
QUIET – IMPATIENT	0.597**	23.886	0.008
GENTLE – AGGRESSIVE	0.415 NS	16.619	0.083
RELIABLE – UNTRUSTWORTHY	0.393 NS	15.701	0.109
GOOD NATURED – STROPPY	0.503*	20.116	0.028

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Appendix 13

Consistency of experienced observers when rating horses on separate constructs Stable 1

CONSTRUCT	W	χ^2	P
INTERESTED DISINTERESTED	0.405***	65.614	0.000
BOLD – TIMID	0.358***	58.070	0.000
FRIENDLY – UNFRIENDLY	0.367***	59.449	0.000
CONFIDENT – NERVOUS	**0.459*****	**74.372**	**0.000**
BORED – ALERT	0.377***	61.018	0.000
RELAXED – TENSE	**0.599*****	**96.984**	**0.000**
LAID-BACK – SPIRITED	0.369***	59.744	0.000
CALM – EXCITED	0.343***	55.639	0.000
UNCONCERNED – WORRIED	**0.457*****	**74.072**	**0.000**
CONTENT – UNHAPPY	0.276***	44.649	0.000
SOCIABLE – UNSOCIABLE	**0.506*****	**82.044**	**0.000**
ACTIVE – PASSIVE	0.157**	25.498	0.002
SECURE – FRIGHTENED	**0.431*****	**69.857**	**0.000**
SPOOKY – TRUSTWORTHY	0.384***	62.146	0.000
SURE – UNSURE	**0.432*****	**70.010**	**0.000**
QUIET – ENERGETIC	0.223***	36.157	0.000
TRUSTING – UNTRUSTING	0.364***	59.012	0.000
PLACID – HOT TEMPERED	0.276***	44.638	0.000
SAFE – JUMPY	0.353***	57.190	0.000
DOCILE – HIGHLY STRUNG	0.301***	48.700	0.000
SETTLED – UNSETTLED	**0.400*****	**64.727**	**0.000**
PLAYFUL – SERIOUS	0.201***	32.520	0.000
AWARE – OBLIVIOUS	0.189***	30.585	0.000
GENTLE – AGGRESSIVE	0.097NS	15.679	0.074
AFFECTIONATE – UNAFFECTIONATE	**0.442*****	**71.554**	**0.000**
STEADY – FLIGHTY	0.371***	60.035	0.000
GOOD NATURED – DIFFICULT	0.236***	38.203	0.000
INTELLIGENT – STUPID	0.109*	17.619	0.040
EASY GOING – STUBBORN	0.261***	42.287	0.000
EXPERIENCED – INEXPERIENCED	0.358***	57.951	0.000
PATIENT – IMPATIENT	0.225***	36.403	0.000

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Consistency of experienced observers rating horses on separate constructs – Stable 2.

CONSTRUCT	W	χ^2	P
INTERESTED – DISINTERESTED	0.156**	23.705	0.003
BOLD – TIMID	0.370***	63.324	0.000
FRIENDLY – UNFRIENDLY	0.304***	52.018	0.000
CONFIDENT – NERVOUS	0.323***	55.153	0.000
BORED – ALERT	0.450***	76.921	0.000
RELAXED – TENSE	0.545***	93.130	0.000
LAID-BACK – SPIRITED	**0.707*** **	**120.940**	**0.000**
CALM – EXCITED	**0.688*** **	**104.578**	**0.000**
UNCONCERNED – WORRIED	0.498***	85.235	0.000
CONTENT – UNHAPPY	0.356***	60.809	0.000
SOCIABLE – UNSOCIABLE	0.311***	53.236	0.000
ACTIVE – PASSIVE	**0.632*** **	**108.045**	**0.000**
SECURE – FRIGHTENED	0.405***	69.204	0.000
SPOOKY – TRUSTWORTHY	0.477***	81.579	0.000
SURE – UNSURE	0.374***	63.885	0.000
QUIET – ENERGETIC	**0.714*** **	**122.069**	**0.000**
TRUSTING – UNTRUSTING	0.252***	43.085	0.000
PLACID – HOT TEMPERED	**0.613*** **	**104.816**	**0.000**
SAFE – JUMPY	0.588***	100.474	0.000
DOCILE – HIGHLY STRUNG	**0.675*** **	**115.470**	**0.000**
SETTLED – UNSETTLED	**0.608*** **	**103.949**	**0.000**
PLAYFUL – SERIOUS	0.175***	29.854	0.000
AWARE – OBLIVIOUS	0.127**	21.713	0.010
GENTLE – AGGRESSIVE	0.440***	75.214	0.000
AFFECTIONATE – UNAFFECTIONATE	0.176***	30.127	0.000
STEADY – FLIGHTY	**0.708*** **	**121.094**	**0.000**
GOOD NATURED – DIFFICULT	0.340***	58.164	0.000
INTELLIGENT – STUPID	0.041NS	7.030	0.634
EASY GOING – STUBBORN	0.315***	53.892	0.000
EXPERIENCED – INEXPERIENCED	0.206***	35.239	0.000
PATIENT – IMPATIENT	0.537***	91.889	0.000

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Appendix 14

Consistency of inexperienced observers rating horses on separate constructs – Stable 1.

CONSTRUCT	W	χ^2	P
INTERESTED – DISINTERESTED	0.427***	**65.326**	**0.000**
FRIENDLY – UNFRIENDLY	**0.480*****	**73.381**	**0.000**
CALM – AGITATED	0.197***	30.090	0.000
CONFIDENT – NERVOUS	0.280***	42.830	0.000
HAPPY – UNHAPPY	0.256***	39.150	0.000
PLACID – HIGHLY STRUNG	0.190***	29.087	0.001
PLAYFUL – SERIOUS	0.391***	59.878	0.000
ACTIVE – PASSIVE	0.099NS	15.105	0.088
SURE – UNSURE	0.239***	36.520	0.000
BORED – INTERESTED	**0.426*****	**65.248**	**0.000**
DOCILE – SPIRITED	0.178***	27.227	0.001
ALERT – TIRED	0.089NS	13.566	0.139
INTERACTIVE – DISTANT	**0.510*****	**77.984**	**0.000**
TOLERANT – INTOLERANT	0.278***	42.492	0.000
RELAXED – TENSE	0.228***	34.832	0.000
SECURE – INSECURE	0.331***	50.627	0.000
TRUSTING – DISTRUSTFUL	0.357***	54.620	0.000
EXTROVERTED – INTROVERTED	0.200***	30.550	0.000
APPROACHABLE – STAND-OFFISH	**0.439*****	**67.167**	**0.000**
AFFECTIONATE – UNAFFECTIONATE	**0.468*****	**71.579**	**0.000**
BOLD – TIMID	0.234***	35.761	0.000
ACCEPTING – REJECTING	0.362***	55.345	0.000
SOCIABLE – UNSOCIABLE	**0.534*****	**81.761**	**0.000**
PATIENT – IMPATIENT	0.267***	40.879	0.000
GENTLE – HARSH	0.254***	38.885	0.000
WELL BEHAVED – STUBBORN	0.305***	46.640	0.000

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Consistency of inexperienced observers rating horses on separate constructs – Stable 2

CONSTRUCT	W	χ^2	P
INTERESTED – DISINTERESTED	0.301***	46.049	0.000
FRIENDLY – UNFRIENDLY	0.395***	60.503	0.000
CALM – AGITATED	**0.654*** **	**100.097**	**0.000**
CONFIDENT – NERVOUS	0.340***	52.048	0.000
HAPPY – UNHAPPY	0.305***	46.615	0.000
PLACID – HIGHLY STRUNG	**0.667*** **	**102.024**	**0.000**
PLAYFUL – SERIOUS	0.385***	58.880	0.000
ACTIVE – PASSIVE	**0.717*** **	**109.668**	**0.000**
SURE – UNSURE	0.250***	38.237	0.000
BORED – INTERESTED	0.282***	43.191	0.000
DOCILE – SPIRITED	**0.766*** **	**117.273**	**0.000**
ALERT – TIRED	0.311***	47.546	0.000
INTERACTIVE – DISTANT	0.327***	50.022	0.000
TOLERANT – INTOLERANT	0.425***	65.010	0.000
RELAXED – TENSE	0.469***	71.737	0.000
SECURE – INSECURE	0.345***	52.733	0.000
TRUSTING – DISTRUSTFUL	0.359***	54.925	0.000
EXTROVERTED – INTROVERTED	0.380***	58.121	0.000
APPROACHABLE – STAND-OFFISH	0.303***	45.315	0.000
AFFECTIONATE – UNAFFECTIONATE	0.301***	45.991	0.000
BOLD – TIMID	0.396***	60.645	0.000
ACCEPTING – REJECTING	0.339***	51.799	0.000
SOCIABLE – UNSOCIABLE	0.371***	56.702	0.000
PATIENT – IMPATIENT	**0.539*** **	**82.408**	**0.000**
GENTLE – HARSH	**0.547*** **	**83.679**	**0.000**
WELL BEHAVED – STUBBORN	0.463***	70.877	0.000

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Appendix 15

MDS - 3 DIMENSIONAL MODELS

Although recommendations on sample size would consider a sample size of 10 and 11 horses as too small for three dimensions, we have conducted a three-dimensional MDS and identify dimensions using linear regression.

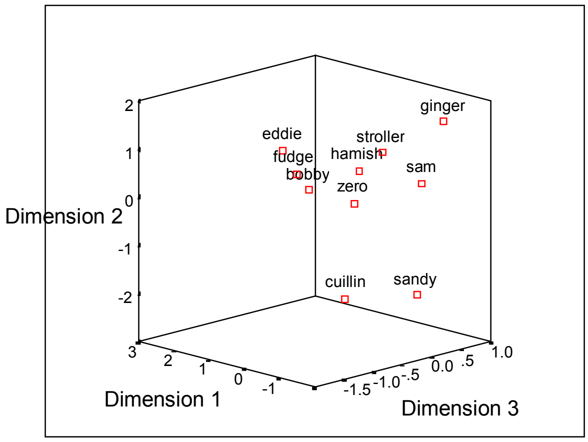
STABLE 1

FAMILIAR OBSERVERS

Euclidean three-dimensional distance model for

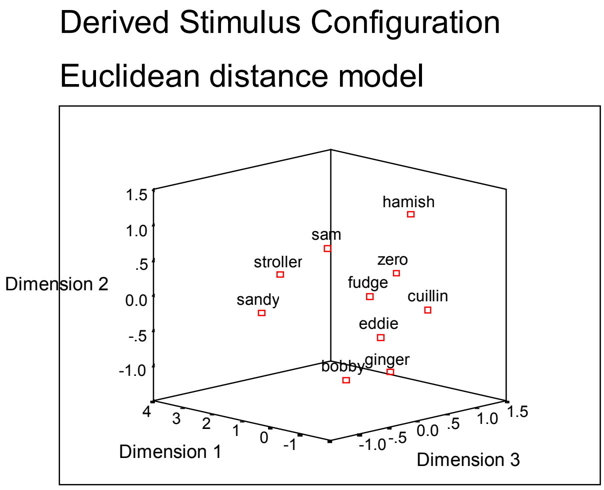
Derived Stimulus Configuration

Euclidean distance model



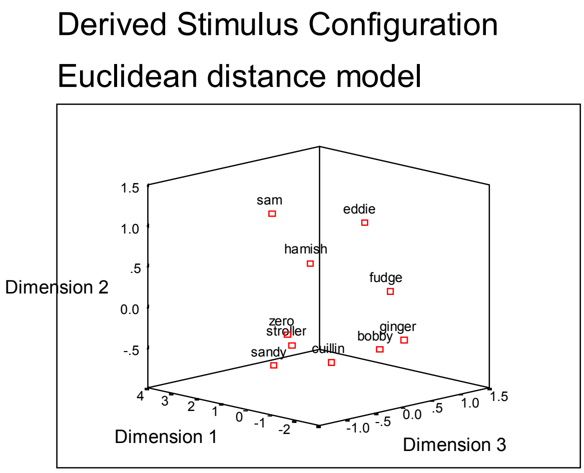
Dimension 1 – Neuroticism (N), Dimension 2 – Extroversion (E), Dimension 3 – Agreeableness (A).

EXPERIENCED OBSERVERS



Dimension 1 (N), Dimension 2 (E), Dimension 3 is unclear.

INEXPERIENCED OBSERVERS



Dimension 1 (E), Dimension 2 (N), Dimension 3 is unclear.

LINEAR REGRESSION – IDENTIFICATION OF DIMENSIONS

STABLE 1

FAMILIAR OBSERVERS

CONSTRUCT	T			Significance		
	DIM 1	DIM 2	DIM 3	DIM 1	DIM 2	DIM 3
1. CONFIDENT-UNSURE	**-13.062*** **	**5.117** **	3.351*	0.000	0.002	0.015
2. HAPPY-UNHAPPY	**-9.085*** **	1.290 NS	-1.292NS	0.000	0.245	0.244

3. LAID-BACK-WORRIER	**-7.191*****	0.660 NS	2.269 NS	0.000	0.534	0.064
4. RELIABLE-UNRELIABLE	-2.896*	-2.575*	-1.980 NS	0.027	0.042	0.095
5. BULLY-BULLIED	-0.363 NS	0.214 NS	0.435 NS	0.729	0.838	0.679
6. INTELLIGENT-STUPID	-2.473*	**3.777****	-1.584 NS	0.048	0.009	0.164
7. WILLING TO WORK-UNWILLING TO WORK	-2.837*	2.299 NS	**-7.694*****	0.030	0.061	0.000
8. STEADY-FAST	**-6.020*****	**-7.889*****	-0.036 NS	0.001	0.000	0.972
9. BOLD-TIMID	**-5.101****	3.359*	2.160 NS	0.002	0.015	0.074
10. FRIENDLY-UNFRIENDLY	-3.952**	1.266 NS	0.655 NS	0.008	0.252	0.537
11. FORWARD GOING-SLOW	1.975 NS	**8.675*****	**-4.612****	0.096	0.000	0.004
12. DOES NOT LIKE SMALL AREAS - NOT WORRIED ABOUT IT	3.050*	-3.073*	**-4.925****	0.023	0.022	0.003
13. PUSHY-DOCILE	-0.084 NS	1.964 NS	1.230 NS	0.936	0.097	0.265
14. TRUSTFUL-SUSPICIOUS	**-9.142*****	-0.489 NS	1.530 NS	0.000	0.642	0.177
15. AFFECTIONATTE-ALOOF	-2.134 NS	3.344*	1.777 NS	0.077	0.016	0.126
16. BOSSY-LIKES TO FOLLOW	-1.839 NS	**4.421****	-0.128 NS	0.116	0.004	0.902
17. OBEDIENT-DISOBEDIENT	-2.091 NS	-3.330*	-1.944 NS	0.082	0.016	0.100
18. EASY TO LEAD-STUBBORN	-5.348**	**-7.600*****	**-4.008****	0.002	0.000	0.007
19. GOOD LEAD HORSE – PREFERS FOLLOWING	-3.998**	4.217**	-1.416 NS	0.007	0.006	0.207
20. INQUISITIVE-NOT INQUISITIVE	-1.066 NS	**10.665*****	1.789 NS	0.327	0.000	0.124
21. RESPONSIVE-UNRESPONSIVE	-0.432 NS	3.170*	**-3.343***	0.681	0.019	0.016
22. ACCEPTING-NOT ACCEPTING	**-6.346*****	-3.613**	0.484 NS	0.001	0.011	0.646
23. SOCIABLE-LONER	-0.925 NS	1.767 NS	0.089 NS	0.391	0.128	0.932
24. LIVELY-SLUGGISH	2.423 NS	**7.525*****	-3.195*	0.052	0.000	0.019
25. PLAYFUL-DULL	2.292 NS	5.416**	0.491 NS	0.062	0.002	0.641
26. QUIET-EXCITABLE	-3.695**	-3.912**	-0.431 NS	0.010	0.008	0.682
27. PATIENT-IMPATIENT	**-6.474*****	-4.257**	-2.400 NS	0.001	0.005	0.053
28. GOOD NATURED-STROPPY	**-5.782*****	0.367 NS	0.299 NS	0.001	0.726	0.775
29. RELAXED-UPTIGHT	**-9.245*****	0.024 NS	**4.047****	0.000	0.982	0.007
30. ALERT-APATHETIC	-0.673 NS	2.828*	-0.718 NS	0.526	0.030	0.500
31. SENSITIVE-UNFEELING	-0.198 NS	1.731 NS	-1.421 NS	0.850	0.134	0.205
32. GREEDY-NOT GREEDY	-2.040 NS	1.502 NS	1.189 NS	0.087	0.184	0.279

Similar to the results derived from the two dimensional model, the first dimension from Stable 1 is associated with the constructs "confident-unsure", "happy-unhappy", "laid-back-worrier", "bold-timid", "trustful-suspicious" and "relaxed-uptight". They can be interpreted as **neuroticism facets**. The second dimension associates with constructs "intelligent-stupid", "steady-fast", "forward going-slow", "easy to lead-stubborn", "inquisitive-not inquisitive" and "lively-sluggish". They can be interpreted as **extroversion facets**. The third dimension associates with "willing to work-unwilling to work" and "responsive-unresponsive" and could be defined as **agreeableness facets**.

EXPERIENCED OBSERVERS

CONSTRUCT	T			Significance		
	DIM 1 (N)	DIM 2 (E)	DIM 3 (Excitability?)	DIM 1	DIM 2	DIM 3

1. INTERESTED-DISINTERESTED	-5.139**	**-6.564*** **	1.042	0.002	0.001	0.337
2. BOLD-TIMID	**6.665*** **	-1.195	0.992	0.001	0.227	0.359
3. FRIENDLY-UNFRIENDLY	-1.016	**-4.870** **	-0.241	0.349	0.003	0.818
4. CONFIDENT-NERVOUS	**13.685*** **	-0.572	0.959	0.000	0.588	0.357
5. BORED-ALERT	**6.893*** **	4.006**	-1.559	0.000	0.007	0.170
6. RELAXED-TENSE	**37.669*** **	-1.124	**-2.890* **	0.000	0.304	0.028
7. LAID BACK-SPIRITED	**10.680*** **	-1.488	-0.570	0.000	0.187	0.589
8. CALM-EXCITED	**9.001*** **	-0.806	-1.906	0.000	0.451	0.105
9. UNCONCERNED-WORRIED	**8.236*** **	0.439	-0.679	0.000	0.676	0.523
10. CONTENT-UNHAPPY	3.100*	-2.178	0.142	0.021	0.072	0.892
11. SOCIABLE-UNSOCIABLE	-2.416	**-12.924*** **	**-2.801* **	0.052	0.000	0.031
12. ACTIVE-PASSIVE	-2.602*	-0.394	-0.586	0.041	0.707	0.579
13. SECURE-FRIGHTENED	**16.659*** **	-3.588**	-0.689	0.000	0.012	0.516
14. SPOOKY-TRUSTWORTHY	**-19.403*** **	2.386	0.141	0.000	0.054	0.893
15. SURE-UNSURE	**16.123*** **	-1.301	-0.620	0.000	0.241	0.558
16. QUIET-ENERGETIC	5.982***	-0.748	-0.374	0.001	0.483	0.721
17. TRUSTING-UNTRUSTING	5.366**	-1.949	-1.952	0.002	0.099	0.099
18. PLACID-HOT TEMPERED	**6.572*** **	-1.932	-0.096	0.001	0.102	0.926
19. SAFE-JUMPY	**11.802*** **	0.089	1.490	0.000	0.932	0.187
20. DOCILE-HIGHLY STRUNG	**7.404*** **	0.084	-0.518	0.000	0.936	0.623
21. SETTLED-UNSETTLED	**9.159*** **	-1.175	-0.759	0.000	0.284	0.477
22. PLAYFUL-SERIOUS	-1.162	**-4.910** **	0.455	0.289	0.003	0.665
23. AWARE-OBLIVIOUS	-5.537***	-3.326*	1.610	0.001	0.016	0.158
24. GENTLE-AGGRESSIVE	0.032	**-2.343**	-0.251	0.975	0.058	0.810
25. AFFECTIONATE-UNAFFECTIONATE	-0.319	**-5.207** **	-0.679	0.761	0.002	0.522
26. STEADY-FLIGHTY	**8.160*** **	-1.845	0.347	0.000	0.115	0.740
27. GOOD NATURED-DIFFICULT	1.735	**-4.102** **	-0.147	0.133	0.006	0.888
28. INTELLIGENT-STUPID	1.795	-2.308	1.164	0.123	0.060	0.289
29. EASY GOING-STUBBORN	4.056**	-1.357	1.970	0.007	0.224	0.096
30. EXPERIENCED-INEXPERIENCED	**7.432*** **	-1.610	-2.257	0.000	0.159	0.065
31. PATIENT-IMPATIENT	3.481*	-3.348*	-1.066	0.013	0.015	0.327

Similar to the results derived from the two dimensional model, the first dimension from the RSD produced by the experts associates with the constructs "confident-nervous", "relaxed-tense", and "secure-frightened", and are correlated with **neuroticism facets** in accordance with five factor concept of personality (Costa and McCrae, Manual). The second dimension associates significantly with the constructs "friendly-unfriendly" and "sociable-unsociable". These constructs could fall within **extroversion facets**. The third dimension is unclear.

INEXPERIENCED OBSERVERS

CONSTRUCT	T			Significance		
	DIM 1 (E)	DIM 2 (N)	DIM 3 (N/Act)	DIM 1	DIM 2	DIM 3
1. INTERESTED-DISINTERESTED	**-6.583***	-0.414	0.934	0.001	0.693	0.386
2. FRIENDLY-UNFRIENDLY	**-10.573***	-0.640	0.351	0.000	0.546	0.738
3. CALM-AGITATED	-3.508*	**4.238**	-2.642*	0.013	0.005	0.038
4. CONFIDENT-NERVOUS	-5.282**	**6.715***	**-5.506**	0.002	0.001	0.002
5. HAPPY-UNHAPPY	-9.470***	0.469	-1.1611	0.000	0.656	0.158
6. PLACID-HIGHLY STRUNG	-4.529**	**2.446***	0.474	0.004	0.050	0.653
7. PLAYFUL-SERIOUS	**-7.506***	-0.585	-1.968	0.000	0.580	0.097
8. ACTIVE-PASSIVE	-1.661	-0.982	-0.994	0.148	0.364	0.358
9. SURE-UNSURE	-4.152**	**3.404***	**-3.100***	0.006	0.014	0.021
10. BORED-INTERESTED	**11.994***	1.318	-0.488	0.000	0.236	0.643
11. DOCILE-SPIRITED	-0.728	**2.493***	1.070	0.494	0.047	0.326
12. ALERT-TIRED	-2.394	0.411	0.846	0.054	0.695	0.430
13. INTERACTIVE-DISTANT	**-11.099***	-1.040	0.148	0.000	0.339	0.887
14. TOLERANT-INTOLERANT	**-7.525***	2.007	-0.244	0.000	0.092	0.815
15. RELAXED-TENSE	-4.783**	2.365	-1.865	0.003	0.056	0.111
16. SECURE-INSECURE	-4.887**	**2.796**	-2.249	0.003	0.031	0.066
17. TRUSTING-DISTRUSTFUL	**-12.666***	2.381	-2.115	0.000	0.055	0.079
18. EXTROVERTED –INTROVERTED	**-18.126***	1.757	**-8.464***	0.000	0.129	0.000
19. APPROACHABLE –STAND-OFFISH	**-15.613***	-1.680	0.208	0.000	0.144	0.842
20. AFFECTIONATE-UNAFFECTIONATE	**-19.659***	-1.933	1.102	0.000	0.101	0.313
21. BOLD-TIMID	-3.012*	-1.370	**-2.459***	0.024	0.220	0.049
22. ACCEPTING-REJECTING	**-9.661***	-0.797	1.688	0.000	0.456	0.142
23. SOCIABLE-UNSOCIABLE	**-15.463***	-1.155	0.814	0.000	0.292	0.447
24. PATIENT-IMPATIENT	-5.352**	**2.821***	0.206	0.002	0.030	0.844
25. GENTLE-HARSH	-4.813**	0.009	**2.715***	0.003	0.993	0.035
26. WELL BEHAVED-STUBBORN	**-6.368***	0.551	0.757	0.001	0.602	0.478

Similarly to the results derived from the two dimensional model, the first dimension from the RSD produced by the naïve raters associates with the constructs "trusting-distrustful", "friendly-unfriendly", "affectionate-unaffectionate", "extroverted-introverted" and "affectionate-unaffectionate". These constructs could fall within **extroversion facets**. The second dimension associates significantly with the constructs "calm-agitated", "confident-nervous", and "docile-spirited", and are correlated with **neuroticism facets.** The third dimension is correlated with "confident-nervous" and "extroverted-introverted" and it is unclear.

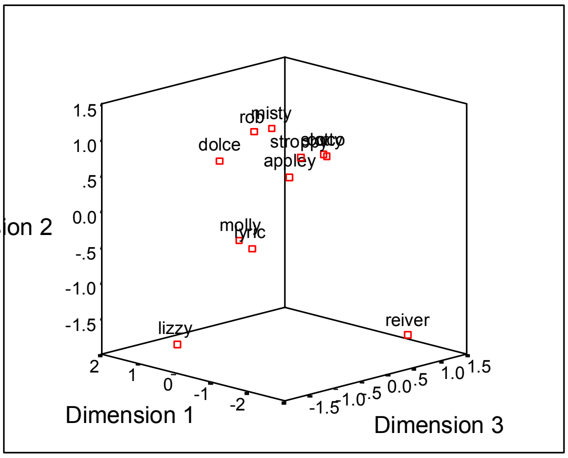
MDS - 3 DIMENSIONAL MODELS

STABLE 2

FAMILIAR OBSERVERS

Derived Stimulus

Euclidean distance

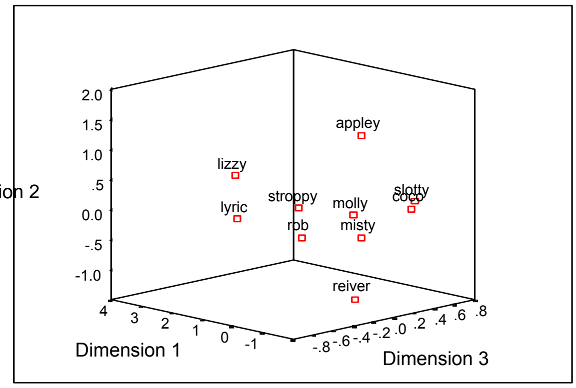


Dimension 1 (N), Dimension 2 (A), Dimension 3 (E).

EXPERIENCED OBSERVERS

Derived Stimulus Configuration

Euclidean distance model

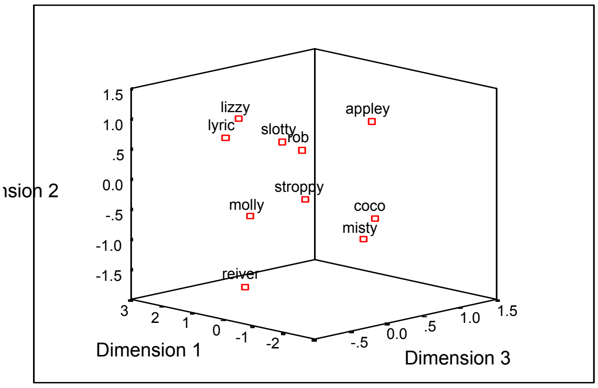


Dimension 1 (E), Dimension 2 (N), Dimension 3 is unclear.

INEXPERIENCED OBSERVERS

Derived Stimulus Configuration

Euclidean distance model



Dimension 1 (N-impulsivness), Dimension 2 (E), Dimension 3 is unclear.

LINEAR REGRESSION – IDENTIFICATION OF DIMENSIONS

STABLE 2

FAMILIAR OBSERVERS

CONSTRUCT	T			Significance		
	DIM 1	DIM 2	DIM 3	DIM 1	DIM 2	DIM 3
1. BOLD – SHY	**5.309*** **	-2.052	3.384*	0.001	0.079	0.012
2. CLEVER – STUPID	2.716*	-3.100*	-0.519	0.030	0.017	0.620
3. SOCIABLE - ANTI-SOCIAL	1.307	**-8.984*** **	**11.793*** **	0.232	0.000	0.000
4. CHEEKY – TIMID	4.633**	-1.485	2.606*	0.002	0.181	0.035
5. FRIENDLY – UNFRIENDLY	-2.207	-3.558**	**5.100*** **	0.063	0.009	0.001
6. CALM – EXCITABLE	**4.602** **	-1.521	-3.227*	0.002	0.172	0.015
7. LAID-BACK – HIGHLY STRUNG	**4.694** **	-1.955	-3.192*	0.002	0.092	0.015
8. AFFECTIONATE – ALOOF	-1.601	-1.810	2.904	0.153	0.113	0.023
9. CONFIDENT – NERVOUS	**10.060*** **	-3.168*	**6.860*** **	0.000	0.016	0.000
10. WILLING TO PLEASE – STUBBORN	-3.949**	**-4.683** **	0.771	0.006	0.002	0.466
11. SAFE – SPOOKY	**6.319*** **	-2.276	-0.249	0.000	0.057	0.811
12. CONTENT – UNSETTLED	**5.919*** **	-4.111**	1.726	0.001	0.005	0.128
13. OBEDIENT – DISOBEDIENT	-2.535*	**-4.053** **	-1.727	0.039	0.005	0.128
14. INTERESTED – DISINTERESTED	-2.916*	**-4.388** **	2.994*	0.022	0.003	0.020
15. LIVELY – DULL	**-4.418** **	-0.409	3.073*	0.003	0.695	0.018
16. EASY TO WORK WITH – DIFFICULT	-2.232	**-6.612*** **	-0.858	0.061	0.000	0.419
17. PLAYFUL – BORING	1.411	-2.100	2.584*	0.201	0.074	0.036
18. SECURE – INSECURE	**6.358*** **	-4.063**	3.281*	0.000	0.005	0.013
19. PREDICTABLE – UNPREDICTABLE	4.044**	-3.618**	-0.111	0.005	0.009	0.915
20. HAPPY – UNHAPPY	3.051*	**-6.550*** **	1.553	0.019	0.000	0.164
21. WELL BEHAVED – NAUGHTY	-3.664**	**-9.211*** **	-3.779**	0.008	0.000	0.007
22. SENSITIVE – UNRESPONSIVE	-3.544**	-2.721*	0.833	0.009	0.030	0.432
23. FRESH – LAZY	**-4.375** **	-1.011	2.735*	0.003	0.346	0.029
24. WELL MANNERED – IGNORANT	-3.723**	**-6.812*** **	-1.980	0.007	0.000	0.088
25. QUIET – IMPATIENT	-0.242	-2.437*	-2.664*	0.816	0.045	0.032
26. GENTLE – AGGRESSIVE	-2.300	-1.758	0.566	0.055	0.122	0.589

27. RELIABLE – UNTRUSTWORTHY	3.877**	**-4.877****	-3.161*	0.006	0.002	0.016
28. GOOD NATURED – STROPPY	**-3.592****	**-3.575****	0.998	0.009	0.009	0.352

The three dimensional model derived from the data from RS indicate that the first dimension associates with the constructs "bold-shy", "calm-excitable", "laid-back-highly strung", "confident-nervous", "safe-spooky" and "secure-insecure". They can be associated with **neuroticism facets**. The second dimension associates with "willing to please-stubborn", "obedient-disobedient", "easy to work with-difficult", "happy-unhappy", "well behaved-naughty" and "well mannered-ignorant". These constructs can be associated with **agreeableness facets**. The third dimension associates with the constructs "sociable-antisocial", "friendly-unfriendly", "confident-nervous" and "playful-boring", and can be associated with **extraversion facets**.

EXPERIENCED OBSERVERS

CONSTRUCT	T			Significance		
	DIM 1 (N)	DIM 2 (E/Act)	DIM 3 (Act)	DIM 1	DIM 2	DIM 3
1. INTERESTED-DISINTERESTED	-0.290	-1.491	**-2.895***	0.782	0.186	0.028
2. BOLD-TIMID	0.295	**7.103*****	0.687	0.778	0.000	0.518
3. FRIENDLY-UNFRIENDLY	2.685*	-0.093	-1.498	0.036	0.929	0.185
4. CONFIDENT-NERVOUS	5.580***	**7.782*****	1.210	0.001	0.000	0.272
5. BORED-ALERT	5.785***	1.891	1.279	0.001	0.108	0.248
6. RELAXED-TENSE	**16.179*****	**4.287****	0.992	0.000	0.005	0.360
7. LAID BACK-SPIRITED	**29.893*****	-0.104	-0.654	0.000	0.920	0.538
8. CALM-EXCITED	**20.026*****	-0.696	-0.802	0.000	0.513	0.453
9. UNCONCERNED-WORRIED	**12.107*****	**3.776****	**2.707***	0.000	0.009	0.035
10. CONTENT-UNHAPPY	5.084**	1.035	0.404	0.002	0.340	0.700
11. SOCIABLE-UNSOCIABLE	2.724*	0.509	-1.890	0.034	0.629	0.108
12. ACTIVE-PASSIVE	**-10.194*****	**5.910*****	-1.467	0.000	0.001	0.193
13. SECURE-FRIGHTENED	**11.053*****	**4.155****	1.194	0.000	0.006	0.278
14. SPOOKY-TRUSTWORTHY	**-11.128*****	-0.705	-0.984	0.000	0.507	0.363
15. SURE-UNSURE	**13.557*****	**11.789*****	**4.201****	0.000	0.000	0.006
16. QUIET-ENERGETIC	**18.234*****	**-5.492****	-0.615	0.000	0.002	0.561
17. TRUSTING-UNTRUSTING	**7.140*****	**3.077***	-0.890	0.000	0.022	0.408
18. PLACID-HOT TEMPERED	**21.894*****	**-4.208****	-0.645	0.000	0.006	0.543
19. SAFE-JUMPY	**15.337*****	-0.664	-2.223	0.000	0.532	0.068
20. DOCILE-HIGHLY STRUNG	**33.029*****	**-3.321***	**-2.636***	0.000	0.016	0.039
21. SETTLED-UNSETTLED	**11.173*****	1.103	1.433	0.000	0.312	0.202
22. PLAYFUL-SERIOUS	-0.841	**3.313***	1.176	0.433	0.016	0.284
23. AWARE-OBLIVIOUS	-1.725	-1.581	-0.860	0.135	0.165	0.423
24. GENTLE-AGGRESSIVE	**13.319*****	**-7.374*****	-1.164	0.000	0.000	0.289
25. AFFECTIONATE-UNAFFECTIONATE	2.598*	-0.525	-1.249	0.041	0.619	0.258
26. STEADY-FLIGHTY	**39.102*****	**-2.995***	-2.215	0.000	0.024	0.069
27. GOOD NATURED-DIFFICULT	**11.066*****	-2.367	-1.794	0.000	0.056	0.123
28. INTELLIGENT-STUPID	0.192	-1.728	1.532	0.854	0.135	0.176
29. EASY GOING-STUBBORN	5.830***	-2.410	-0.376	0.001	0.053	0.720
30. EXPERIENCED-INEXPERIENCED	**8.120*****	1.329	-1.311	0.000	0.232	0.238

31. PATIENT-IMPATIENT	**10.456*** **	**-3.536***	-0.980	0.000	0.012	0.365

The three dimensional model derived from the data from RS indicates that the first dimension associates with the constructs "laid-back-spirited", "calm-excited", quiet-energetic", "placid-hot tempered", "docile-highly strung" and "steady-flighty" and can be associated with **extraversion facets.** The second dimension associates with "bold-timid", "confident-nervous" and "sure-unsure". They can be associated with **neuroticism facets**. The third dimension associates with the constructs "interested-disinterested" and "sure-unsure" and it is not very clear.

INEXPERIENCED OBSERVERS

CONSTRUCT	T			Significance		
	DIM 1 (N)	DIM 2 (E)	DIM 3 (E/Act)	DIM 1	DIM 2	DIM 3
1. INTERESTED-DISINTERESTED	1.212	**7.548*** **	**-4.814** **	0.271	0.000	0.003
2. FRIENDLY-UNFRIENDLY	**7.924*** **	**6.509*** **	-3.591*	0.000	0.001	0.011
3. CALM-AGITATED	**22.266*** **	-1.178	**3.513***	0.000	0.283	0.013
4. CONFIDENT-NERVOUS	4.797**	3.655*	1.214	0.003	0.011	0.270
5. HAPPY-UNHAPPY	**6.632*** **	4.752**	-0.044	0.001	0.003	0.967
6. PLACID-HIGHLY STRUNG	**46.959*** **	-3.992**	0.128	0.000	0.007	0.902
7. PLAYFUL-SERIOUS	-0.899	**4.125** **	0.797	0.403	0.006	0.456
8. ACTIVE-PASSIVE	**-9.328*** **	6.236***	1.761	0.000	0.001	0.129
9. SURE-UNSURE	2.614*	**3.775** **	**2.565***	0.040	0.009	0.043
10. BORED-INTERESTED	0.766	**-3.606***	**2.750***	0.473	0.011	0.033
11. DOCILE-SPIRITED	**30.795*** **	**-10.064*** **	0.018	0.000	0.000	0.986
12. ALERT-TIRED	-4.109**	2.169	-1.238	0.006	0.073	0.262
13. INTERACTIVE-DISTANT	0.255	**5.954*** **	-1.266	0.807	0.001	0.252
14. TOLERANT-INTOLERANT	**13.583*** **	4.001**	-0.008	0.000	0.007	0.994
15. RELAXED-TENSE	**15.129*** **	3.350*	2.040	0.000	0.015	0.088
16. SECURE-INSECURE	**11.218*** **	**5.783*** **	**2.905***	0.000	0.001	0.027
17. TRUSTING-DISTRUSTFUL	**10.728*** **	5.393**	-0.641	0.000	0.002	0.545
18. EXTROVERTED –INTROVERTED	-5.181**	**5.680*** **	0.330	0.002	0.001	0.752
19. APPROACHABLE –STAND-OFFISH	**12.161*** **	**9.491*** **	**-3.117***	0.000	0.000	0.021
20. AFFECTIONATE-UNAFFECTIONATE	**6.019*** **	**5.457** **	**-3.491***	0.001	0.002	0.013
21. BOLD-TIMID	-4.952**	**5.529*** **	1.769	0.003	0.001	0.127
22. ACCEPTING-REJECTING	**9.891*** **	**5.866*** **	**-2.822***	0.000	0.001	0.030
23. SOCIABLE-UNSOCIABLE	**6.162*** **	**10.820*** **	**-2.899***	0.001	0.000	0.027
24. PATIENT-IMPATIENT	**12.676*** **	-0.928	0.327	0.000	0.389	0.755
25. GENTLE-HARSH	**19.833*** **	-1.758	**-3.909** **	0.000	0.129	0.008
26. WELL BEHAVED-STUBBORN	**16.815*** **	-1.086	-1.555	0.000	0.319	0.171

The three dimensional model constructs derived from the data from RS indicate that the first dimension associates with constructs "calm-agitated", "placid-highly strung", "docile-spirited", and "gentle-harsh" and could be associated with **neuroticism (impulsiveness) facets.** The second dimension associates with "interested-disinterested", "playful-serious", "interactive-distant", and "sociable-unsociable". They can be associated with **extroversion facets**. The third dimension associates with the constructs "interested-disinterested" and "gentle-harsh" and it is not very clear.

Appendix 16

Full list of constructs provided by all 30 familiar observers – Study 2

1. FRIENDLY	22	UNFRIENDLY	**12**
2. ATTENTION SEEKER	22	INDEPENDENT	**10**
3. CONFIDENT	20	TIMID	**5**
4. BOLD	19	SHY	**13**
5. EASY TO HANDLE	17	DIFFICULT	**9**
6. SOCIABLE	16	UNSOCIABLE	**13**
7. BOSSY	16	SUBMISSIVE	**10**
8. HAPPY	16	UNHAPPY	**9**
9. LAID BACK	16	HIGHLY STRUNG	**6**
10. CALM	16	EXCITABLE	**5**
11. PLAYFUL	14	BORING	**6**
12. AFFECTIONATE	14	AGGRESSIVE	**5**
13. OBEDIENT	13	DISOBEDIENT	**8**
14. NERVOUS	13	BOLD/CONFIDENT	**4/3**
15. INTERESTED	12	DISINTERESTED	**5**
16. AGGRESSIVE	12	GENTLE FRIENDLY	4 3
17. DOMINANT	11	SUBMISSIVE	5
18. GENTLE	11	ROUGH	4
19. EXPERIENCED	10	INEXPERIENCED	**9**
20. MATURE	10	IMMATURE	**6**
21. EASY TO WORK WITH	10	DIFFICULT STUBBORN	**4** **3**
22. CHEEKY	10	WELL BEHAVED	**3**
23. SECURE	9	INSECURE	**9**
24. NASTY	9	NICE	**7**
25. BRAVE	9	SCARED	**3**
26. FORWARD	9	LAZY HESITANT	3 2
27. BULLY	9	BULLIED	**2**
28. STUBBORN	8	WILLING	5
29. INTELLIGENT	8	THICK STUPID	**3** **3**

30. RELAXED	8	TENSE	**3**
31. CONTENT	8	GRUMPY	**1**
32. PATIENT	7	IMPATIENT	7
33. CLEVER	7	STUPID	4
34. BARGY	7	QUIET	**4**
35. TRUSTWORTHY	6	UNTRUSTWORTHY	6
36. AWARE	6	UNAWARE	**6**
37. GRUMPY	6	HAPPY AFFECTIONATE	**4** **2**
38. QUIET	6	BOISTEROUS	2
39. PUSHY	6	RESPECTFUL/QUIET	**1**
40. PLACID	6	FIERY	**1**
41. SURE	5	UNSURE	**5**
42. FLIGHTY	5	STEADY	4
43. EASY TO CATCH	5	HARD TO CATCH	4
44. FRIGHTENED	5	BRAVE	**2**
45. RESPONSIVE	5	UNRESPONSIVE	**1**
46. JUMPY	5	CALM LAID-BACK STEADY	1 1 1
47. ACTIVE	4	LAZY	4
48. CURIOUS	4	UNINTERESTED	**3**
49. INQUISITIVE	4	UNINTERESTED	**2**
50. LIVELY	4	DULL	2
51. ARROGANT	4	PROTECTIVE KIND	1 1
52. MAKES FACES	4	HAPPY ACCEPTING GOOD-NATURED	1 1 1
53. DETERMINED	4	FLEXIBLE	1
54. TRUSTING	4	DISTRUSTING	1
55. LEADER	3	FOLLOWER	3
56. SENSIBLE	3	SILLY	2
57. STROPPY	3	COMPLIABLE	1
58. POLITE	3	RUDE	1
59. SPOOKY	3	FLIGHTY BOMB-PROOF	1 1
60. STAND OFFISH	3	TIMID APPROACHABLE	1 1
61. RELIABLE	2	UNRELIABLE	2
62. WARY	2	SELF ASSURED	**1**

63. BARGY	2	CONSIDERATE	**1**
64. BOLSHY	2	AMENABLE	1
65. OUTGOING	2	RESERVED	1
66. ENERGETIC	2	PASSIVE LAZY	1 1
67. PREDICTABLE	1	UNPREDICTABLE	3
68. TOLERANT	1	IRRITABLE	1
69. TACTILE	1	REMOTE	1
70. RESPECTFUL	1	DISRESPECTFUL	1
71. CONSIDERATE	1	INCONSIDERATE	1
72. EXCITABLE	1	SLOW	1

Appendix 17

Uniform scoring list for familiar observers – Study 2

FRIENDLY	1	2	3	4	5	UNFRIENDLY
ATTENTION SEEKER	1	2	3	4	5	**INDEPENDENT**
CONFIDENT	1	2	3	4	5	**TIMID**
BOLD	1	2	3	4	5	**SHY**
EASY TO HANDLE	1	2	3	4	5	**DIFFICULT**
SOCIABLE	1	2	3	4	5	**UNSOCIABLE**
BOSSY	1	2	3	4	5	**SUBMISSIVE**
HAPPY	1	2	3	4	5	**UNHAPPY**
LAID BACK	1	2	3	4	5	**HIGHLY STRUNG**
CALM	1	2	3	4	5	**EXCITABLE**
PLAYFUL	1	2	3	4	5	**BORING**
AFFECTIONATE	1	2	3	4	5	**AGGRESSIVE**
OBEDIENT	1	2	3	4	5	**DISOBEDIENT**
NERVOUS	1	2	3	4	5	**BOLD**
INTERESTED	1	2	3	4	5	**DISINTERESTED**
AGGRESSIVE	1	2	3	4	5	**GENTLE**
DOMINANT	1	2	3	4	5	**SUBMISSIVE**
GENTLE	1	2	3	4	5	**ROUGH**
EXPERIENCED	1	2	3	4	5	**INEXPERIENCED**
MATURE	1	2	3	4	5	**IMMATURE**
EASY TO WORK WITH	1	2	3	4	5	**DIFFICULT**
CHEEKY	1	2	3	4	5	**WELL BEHAVED**
SECURE	1	2	3	4	5	**INSECURE**
NASTY	1	2	3	4	5	**NICE**
BRAVE	1	2	3	4	5	**SCARED**
FORWARD	1	2	3	4	5	**HESITANT**
BULLY	1	2	3	4	5	**BULLIED**
STUBBORN	1	2	3	4	5	**WILLING**

INTELLIGENT	1	2	3	4	5	**THICK**
RELAXED	1	2	3	4	5	**TENSE**
CONTENT	1	2	3	4	5	**GRUMPY**
PATIENT	1	2	3	4	5	**IMPATIENT**
CLEVER	1	2	3	4	5	**STUPID**
BARGY	1	2	3	4	5	**QUIET**

Appendix 18
Appendix 9.2

Instructions attached to the scoring forms

HORSE PERSONALITY SCORING FORM

Instructions for use

Dasha Grajfoner, Psychology Department, The University of Edinburgh

Please read the instructions before beginning.
On the top of the scoring sheet please write your name in the space provided.
Each scoring list is for one of the horses you know. There is name of the horse on the top of each page.

This scoring list contains 34 opposite descriptors with a 5 point scoring scale e.g.:

FRIENDLY	1	2	3	4	5	**UNFRIENDLY**

All words were chosen by the majority of you. Please read them carefully and circle one number that best corresponds to the individual horse. When scoring please try to sum up an overall personality of the horse, e.g. how he/she behaves in different situations (stable, outside, riding, handling etc.).

In the above case

Circle **1** if you think the horse is **VERY friendly**

Circle **2** if you think the horse is **FAIRLY friendly**

Circle **3** if you think the horse is **NEITHER friendly nor unfriendly**

Circle **4** if you think the horse is **FAIRLY unfriendly**

Circle **5** if you think the horse is **VERY unfriendly**

Please follow this example when you are assessing the horse on all 34 descriptors. Circle one number in each pair of words. It is important to assess all horses on all 34 pairs of words.

There are no right or wrong answers.
IMPORTANT: DO NOT DISCUSS THE SCORING WITH YOUR COLLEAGUES BEFORE OR BETWEEN THE SCORING!

Should you have any questions before or while completing the questionnaires please give me a call on **0776 55 26 347.**
Thank you! Dasha Grajfoner

Appendix 19

HORSE PERFORMANCE SCORES

The following scoring list gives you an opportunity to assess the horse's overall performance in most situations (how efficient is the horse for riding, how cooperative it is etc). There is a 9 point scoring scale for each horse. Please circle 1 if you think that the horse's overall working performance is very poor and circle 9 if you think that the horse's overall performance is excellent. Choose grades in between according to how do you think the horse performs, e.g. for an average performance you might circle 5. There are no "right" or "wrong" answers and no "good" or "bad" choices.

RSPCA

NAME:
DATE:

PERFORMANCE

	POOR								EXCELLENT
MILLIE	1	2	3	4	5	6	7	8	9
HARRY	1	2	3	4	5	6	7	8	9
JAZZ	1	2	3	4	5	6	7	8	9
FLASH	1	2	3	4	5	6	7	8	9
LIL	1	2	3	4	5	6	7	8	9
QUENTIN	1	2	3	4	5	6	7	8	9
MR BEAN	1	2	3	4	5	6	7	8	9
CANDY	1	2	3	4	5	6	7	8	9
KURT	1	2	3	4	5	6	7	8	9
MOLLY	1	2	3	4	5	6	7	8	9
ROCKY	1	2	3	4	5	6	7	8	9
WILLOW	1	2	3	4	5	6	7	8	9

BLOSSOM	1	2	3	4	5	6	7	8	9

Appendix 20

Scoring list produced by unfamiliar observers – Study 2

INTERESTED	1	2	3	4	5	UNINTERESTED
BOLD	1	2	3	4	5	SHY
CURIOUS	1	2	3	4	5	INDIFFERENT
PLAYFUL	1	2	3	4	5	BORED
OUTGOING	1	2	3	4	5	TIMID
CALM	1	2	3	4	5	NERVOUS
HAPPY	1	2	3	4	5	SAD
ENTERTAINED	1	2	3	4	5	BORED
AT EASE	1	2	3	4	5	SCARED
INQUISITIVE	1	2	3	4	5	DISINTERESTED
CONFIDENT	1	2	3	4	5	WARY
FRIENDLY	1	2	3	4	5	UNFRIENDLY
BRAVE	1	2	3	4	5	SCARED
LAID-BACK	1	2	3	4	5	WORRIED
CALM	1	2	3	4	5	BOISTEROUS
AFFECTIONATE	1	2	3	4	5	COLD
STEADY	1	2	3	4	5	JUMPY
NOSY	1	2	3	4	5	NOT BOTHERED
STRONG	1	2	3	4	5	ANXIOUS
DOMINANT	1	2	3	4	5	SUBMISSIVE
ENERGETIC	1	2	3	4	5	LAZY
RELAXED	1	2	3	4	5	UPTIGHT
ALERT	1	2	3	4	5	UNALERT
GENTLE	1	2	3	4	5	HARD
COMFORTABLE	1	2	3	4	5	UNCOMFORTABLE
QUIET	1	2	3	4	5	LOUD
EASY GOING	1	2	3	4	5	STUBBORN

SLOW	1	2	3	4	5	**FAST**

Appendix 21

Kendall's coefficient of concordance W – all constructs – Stable 1

HORSE	W - FAMILIAR	W- UNFAMILIAR
1. RYAN	0.458***	0.340***
2. JACK	0.643***	0.338***
3. BARNEY	0.643***	0.313***
4. DICE	0.308***	0.229***
5. BLUE	0.559***	0.273***
6. TIA	0.473***	0.145 NS
7. RHUM	0.615***	0.291***
8. DANNY	0.415***	0.246***
9. DIZZY	0.395***	0.171*
10. ZAK	0.707***	0.482***
11. TICA	0.664***	0.398***
12. GUNTHER	0.768***	0.343***
13. DUKE	0.552***	0.346***

Raters scoring horses on all constructs – Stable 2

HORSE	W - FAMILIAR	W- UNFAMILIAR
1. MILLIE	0.476***	0.322***
2. HARRY	0.463***	0.366***
3. JAZZ	0.798***	0.366***
4. FLASH	0.727***	0.334***
5. LIL	0.413***	0.214***
6. QUENTIN	0.588***	0.447***
7. MR. BEAN	0.476***	0.343***
8. CANDY	0.391***	0.280***
9. KURT	0.485***	0.491***
10. MOLLY	0.516***	0.537***
11. ROCKY	0.498***	0.230***
12. WILLOW	0.502***	0.247***
13. BLOSSOM	0.658***	0.181**

Raters scoring horses on all constructs – Stable 2
(familiar observers only)

HORSE	W – FAMILIAR
1. BRUNO	0.346***

2. SOPHIE	0.562***
3. FLEUR	0.534***
4. GANDALF	0.681***
5. JESTER	0.426***
6. DILLY	0.427***
7. CORK	0.215***
8. PETRA	0.373***
9. DANNY	0.623***
10. CUILLEN	0.588***
11. PANDORA	0.622***
12. BESS	0.283***

Appendix 22

The difference between elicited and partially provided constructs for Stable 1

HORSE	Familiar W		Unfamiliar W	
	Majority	Provided	Majority	Provided
RYAN	.395***	.307***	.263***	.345***
JACK	.574***	.624***	.403***	.274***
BARNEY	.523***	.658***	.322***	.268***
DICE	.292**	.152NS	.226**	.265***
BLUE	.565***	.423***	.221***	.384***
TIA	.316***	.434***	.151NS	.176*
RHUM	.415***	.622***	.255**	.319***
DANNY	.308***	.384***	.203*	.280***
DIZZY	.263**	.289**	.122NS	.150NS
ZAK	.686***	.594***	.256**	.592***
TICA	.501***	.714***	.138NS	.474***
GUNTHER	.674***	.791***	.174*	.485***
DUKE	.534***	.462***	.172*	.402***
Mean Kendall's W	0.465	0.496	0.223	0.339

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

The difference between elicited and partially provided constructs for Stable 2

HORSE	Familiar W		Unfamiliar W	
	Majority	Provided	Majority	Provided
MILLIE	.518***	.394***	.223*	.403***
HARRY	.417***	.415***	.331***	.340***
JAZZ	.790***	.769***	.437***	.290***
FLASH	.615***	.762***	.309***	.378***
LIL	.492***	.373***	.196*	.200*
QUENTIN	.601***	.413***	.446***	.385***
MR. BEAN	.430***	.430***	.097NS	.519***
CANDY	.412***	.273***	.134NS	.306***
KURT	.585***	.419***	.361***	.549***
MOLLY	.529***	.433***	.468***	.453***
ROCKY	.421***	.450***	.148NS	.315***
WILLOW	.433***	.454***	.154NS	.322***
BLOSSOM	.479***	.734***	.102NS	.262***
Mean Kendall's W	0.517	0.486	0.262	0.363

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

The difference between elicited and partially provided constructs for Stable 3

HORSE	Familiar W	
	Majority	Provided
BRUNO	.287***	.267***
SOPHIE	.536***	.587***
FLEUR	.531***	.563***
GANDALF	.634***	.671***
JESTER	.321***	.429***
DILLY	.459***	.358***
CORK	.133NS	.130NS
PETRA	.361***	.358***
DANNY	.654***	.529***
CUILLEN	.574***	.525***
PANDORA	.697***	.555***
BESS	.174NS	.269***
Mean Kendall's W	0.446	0.436

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

Appendix 23

Difference between constructs elicited by at least 2/3 of the participants and constructs elicited by 1/3 of the participants

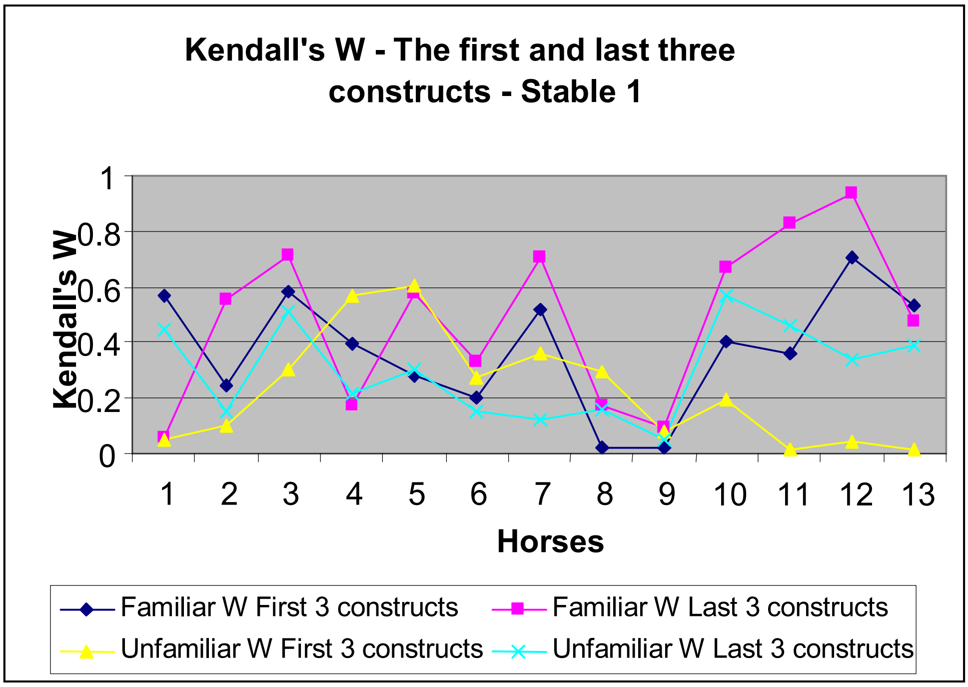
Stable 1

We looked at what the differences were in the degree of agreement between the three constructs elicited by 2/3 of the familiar observers (at least 20 participants) and three constructs elicited by 1/4 of the familiar participants (7).
We also looked at the degree of agreement when unfamiliar observers rated horses on the first and last three constructs on the scoring list.

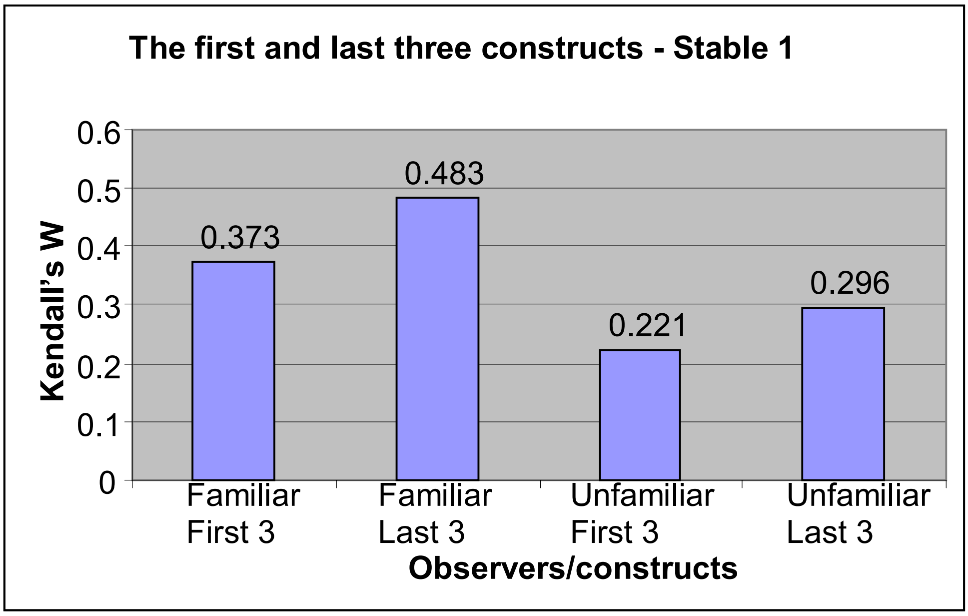
Kendall's W for the first and last 3 constructs – both observer groups – Stable1

HORSE	Familiar W		Unfamiliar W	
	First 3 constructs	Last 3 constructs	First 3 constructs	Last 3 constructs
RYAN	.568*	.055NS	.052NS	.443*
JACK	.244NS	.551*	.100NS	.154NS
BARNEY	.581**	.712**	.300*	.509**
DICE	.397*	.176NS	.565**	.217NS
BLUE	.283NS	.573**	.604**	.304*
TIA	.203NS	.333NS	.271*	.148NS
RHUM	.519*	.707**	.357*	.120NS
DANNY	.025NS	.173NS	.292*	.156NS
DIZZY	.025NS	.091NS	.076NS	.048NS
ZAK	.404*	.672**	.195NS	.565**
TICA	.362*	.827***	.016NS	.459**
GUNTHER	.707**	.938***	.041NS	.338*
DUKE	.534*	.478*	.014NS	.388*
Mean Kendall's W	0.371	0.483	0.221	0.296

*** P<0.001; **P<0.01; *P<0.05; NS – not significant



Mean Kendall's W for the first and last 3 constructs – both observer groups – Stable 1



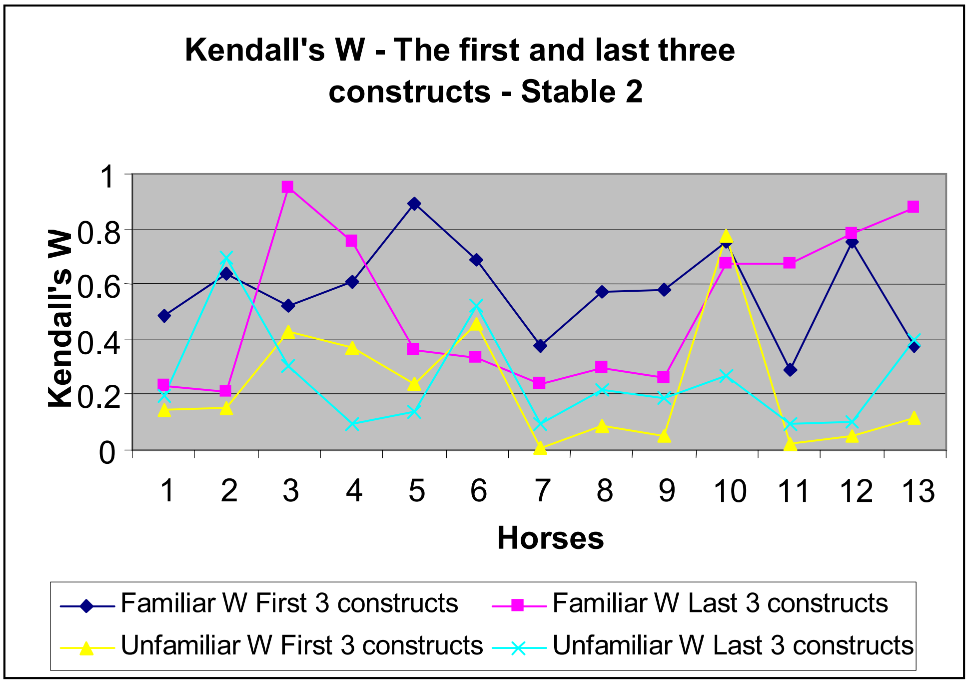
Kendall's W for the first and last 3 constructs – both observer groups – Stable 1

The mean values show a tendency for higher agreement when horses were rated on constructs provided by only three observers. However, the differences are not significant either for familiar (t=1.581, p>0.05) or unfamiliar observers (t=0.932, p>0.05).

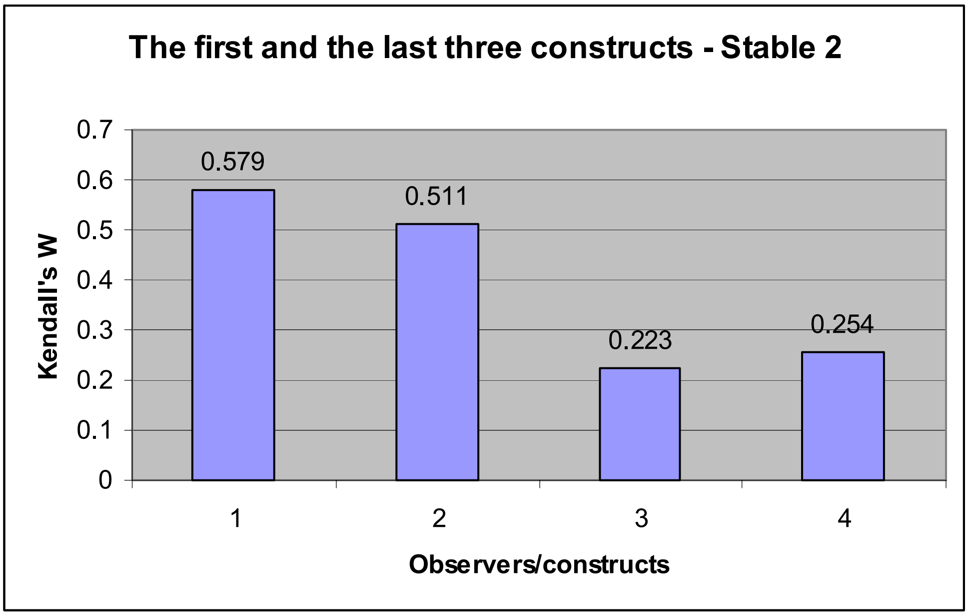
Kendall's W for the first and last 3 constructs – both observer groups – Stable 2

HORSE	**Familiar W**		**Unfamiliar W**	
	First 3 constructs	Last 3 constructs	First 3 constructs	Last 3 constructs
MILLIE	.483*	.232 NS	.144 NS	.196 NS
HARRY	.636**	.212NS	.150 NS	.696***
JAZZ	.521*	.950***	.429*	.305*
FLASH	.609**	.755**	.370*	.097 NS
LIL	.894***	.363 NS	.242 NS	.137 NS
QUENTIN	.692**	.333 NS	.458**	.520**
MR. BEAN	.375*	.236 NS	.009 NS	.091 NS
CANDY	.569*	.297 NS	.090 NS	.217 NS
KURT	.578*	.264 NS	.050 NS	.191 NS
MOLLY	.755**	.674**	.772***	.268 NS
ROCKY	.292 NS	.672**	.025 NS	.091 NS
WILLOW	.755**	.782**	.053 NS	.105 NS
BLOSSOM	.380*	.875***	.117 NS	.397*
Mean Kendall's W				

*** P<0.001; **P<0.01; *P<0.05; NS – not significant



Mean Kendall's W for the first and last 3 constructs – both observer groups – Stable 2



Mean Kendall's W for the first and last 3 constructs – both observer groups – Stable 2

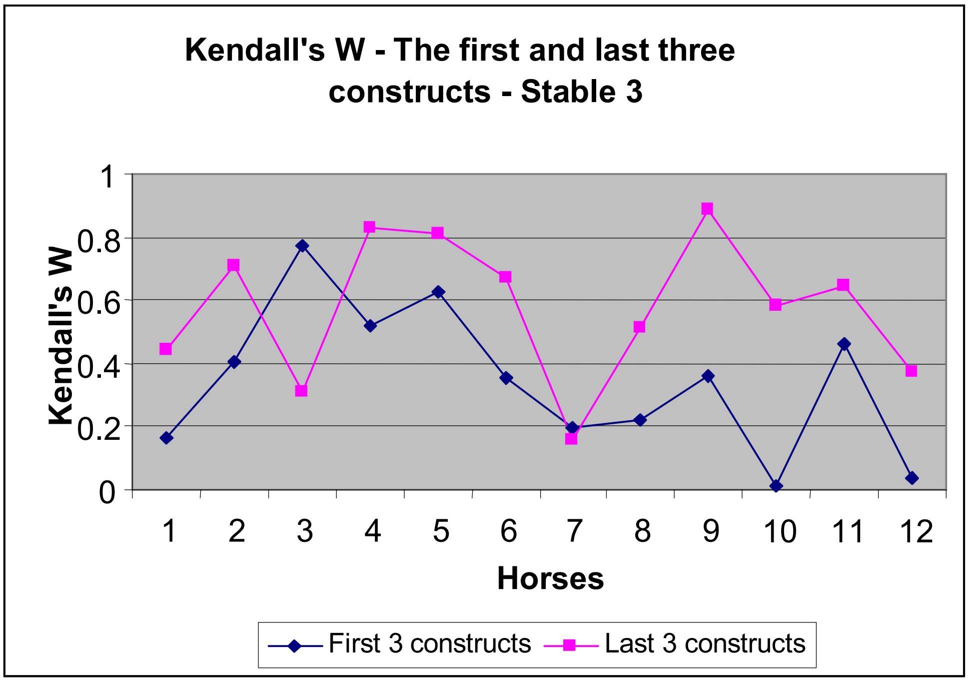
There was no particular pattern in the second stable either. The familiar observers agreed more on the first three constructs, the unfamiliar on the last three. T-test showed no significant difference either in the group of familiar observers ($t=0.73$, $p>0.05$) or in the group of unfamiliar observers ($t=0.44$, $p>0.05$).

Kendall's W for the first and last 3 constructs – Stable 3

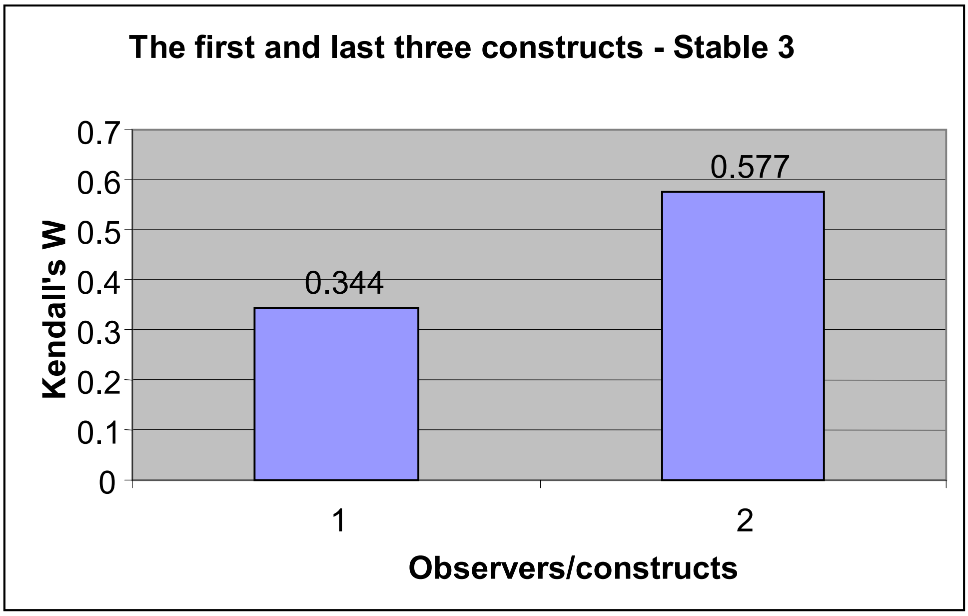
HORSE	**Familiar W**	
	First 3 constructs	Last 3 constructs
BRUNO	.164 NS	.444*
SOPHIE	.406*	.710**
FLEUR	.770***	.309 NS
GANDALF	.520**	.831***
JESTER	.626**	.808***
DILLY	.352*	.670**
CORK	.199 NS	.157 NS
PETRA	.222 NS	.513**
DANNY	.361*	.885***
CUILLEN	.011 NS	.583**
PANDORA	.462*	.646**
BESS	.039 NS	.374*
Mean Kendall's'W	0.344	0.577

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

There was a significant difference in mean values in Kendall's W when familiar observers rated the horses on the first three and on the last three constructs ($t=3.002$, $p<0.05$). The observers reached higher agreement when they used descriptors provided by the minority.



Mean Kendall's W for the first and the last 3 constructs – both observer groups – Stable 3



Mean Kendall's W for the first and last 3 constructs – both observer groups – Stable 3

Appendix 24

Linear regression to identify two dimensions – all stables – Familiar observers

CONSTRUCTS		T – Regression coefficient					
		Stable 1		Stable 2		Stable 3	
		DIM 1	DIM 2	DIM 1	DIM 2	DIM 1	DIM2
FRIENDLY	UNFRIENDLY	**6.298*****	2.484*	-1.040 NS	2.812*	-2.935*	-.842 NS
ATTENTION SEEKER	INDEPENDENT	-1.192 NS	**2.769***	-.987 NS	**3.143****	**-3.754****	-.284 NS
CONFIDENT	TIMID	-1.921 NS	.184 NS	**-7.766*****	-2.558*	**-11.464*****	-.031 NS
BOLD	SHY	-2.920*	.169 NS	**-8.539*****	-2.924*	**-14.725*****	**3.750****
EASY TO HANDLE	DIFFICULT	**13.354*****	1.211 NS	**-5.685*****	1.266 NS	-.898 NS	**-3.722****
SOCIABLE	UNSOCIABLE	**7.198*****	**3.794****	-2.691*	**4.563*****	-1.352 NS	-.017 NS
BOSSY	SUBMISSIVE	-3.778**	-.229 NS	-3.359**	-1.303 NS	-2.539*	.651 NS
HAPPY	UNHAPPY	2.308*	2.106 NS	-2.511*	4.140**	-4.101**	-1.377 NS
LAID BACK	HIGHLY STRUNG	**7.335*****	-2.772*	-2.682*	.791 NS	-1.466 NS	-2.600*
CALM	EXCITABLE	**6.445*****	**-5.147*****	-2.110 NS	.560 NS	-.896 NS	-2.231 NS
PLAYFUL	BORING	-3.498**	3.812**	-1.078 NS	**2.585***	-3.158*	.904 NS
AFFECTIONATE	AGGRESSIVE	4.138**	2.858*	.042 NS	**4.837*****	-.333 NS	-1.084 NS
OBEDIENT	DISOBEDIENT	**5.780*****	-.078 NS	-2.835*	2.348*	1.078 NS	**-3.872****
NERVOUS	BOLD	1.037NS	1.271 NS	**8.047*****	3.190**	**13.053*****	1.415 NS
INTERESTED	DISINTERESTE D	**-5.533*****	**5.268*****	-1.334 NS	**4.595*****	-2.646*	.782 NS
AGGRESSIVE	GENTLE	**-10.528*****	**-3.457****	.270 NS	**-3.614****	-1.937 NS	2.164 NS
DOMINANT	SUBMISSIVE	**-5.000*****	-.619 NS	-2.440*	-2.017 NS	-2.243 NS	2.023 NS
GENTLE	ROUGH	**12.818*****	**3.154****	.333 NS	**4.144****	**5.356*****	**-5.136*****
EXPERIENCED	INEXPERIENCE D	1.320 NS	.760 NS	**-7.674*****	-3.803**	-2.784*	-2.219 NS
MATURE	IMMATURE	1.838NS	-.216 NS	**-5.619*****	-3.047*	-1.835 NS	-3.035*
EASY TO WORK WITH	DIFFICULT	**14.564*****	.644 NS	-3.872**	1.845 NS	-1.074 NS	**-5.439*****
CHEEKY	WELL BEHAVED	**-8.174*****	.287 NS	.106 NS	.488 NS	-3.156*	1.001 NS
SECURE	INSECURE	1.824NS	-.029 NS	**-12.126*****	.926 NS	**-9.807*****	**-4.552*****
NASTY	NICE	**-8.556*****	-3.111*	-.067 NS	**-4.565*****	-2.403*	2.805*
BRAVE	SCARED	-1.331NS	-.410 NS	**-8.922*****	-.324 NS	**-6.990*****	-.253 NS
FORWARD	HESITANT	**-5.035*****	.716 NS	**-6.971*****	.165 NS	**-5.339*****	.919 NS
BULLY	BULLIED	-3.482**	-1.332 NS	-2.559*	.218 NS	-1.575 NS	.966 NS
STUBBORN	WILLING	-1.567NS	-1.839 NS	1.017 NS	**-4.197****	-.847 NS	2.264*
INTELLIGENT	THICK	-2.697*	2.630*	-2.716*	-.290 NS	-1.816 NS	1.514 NS
RELAXED	TENSE	**4.742*****	-3.129*	**-4.983*****	.282 NS	**-8.162*****	**-4.574*****
CONTENT	GRUMPY	**6.342*****	.064 NS	-3.096**	**5.720*****	-2.387*	-.544 NS
PATIENT	IMPATIENT	**8.689*****	-3.863**	-1.697 NS	1.381 NS	.990 NS	**-4.620*****
CLEVER	STUPID	-2.414*	3.065*	**-4.792*****	.313 NS	-3.469**	1.230 NS

BARGY	QUIET	**-10.382*** **	1.002 NS	-.025 NS	-1.557 NS	-4.348**	3.655**

P<0.05; ** P<0.01; *** P<0.001; NS - not significant.

Linear regression to identify two dimensions - Stable 1 and 2 – Unfamiliar observers

CONSTRUCTS		T – Regression coefficient			
		Stable 1		Stable 2	
		DIM 1	DIM 2	DIM 1	DIM 2
INTERESTED	UNINTERESTED	2.401*	6.574***	10.923***	-2.230*
BOLD	SHY	7.688***	2.097 NS	10.651***	3.596**
CURIOUS	INDIFFERENT	1.575 NS	4.437***	7.430***	-3.502**
PLAYFUL	BORED	5.355***	2.300*	11.297***	-.367 NS
OUTGOING	TIMID	7.265***	.952 NS	9.202***	2.956*
CALM	NERVOUS	2.674*	-.154 NS	1.862 NS	.527 NS
HAPPY	SAD	5.750***	1.741 NS	6.697***	-.472 NS
ENTERTAINED	BORED	3.385**	2.072 NS	4.925***	-1.979 NS
AT EASE	SCARED	3.490**	.775 NS	1.996 NS	.671 NS
INQUISITIVE	DISINTERESTED	3.003*	3.299**	10.615***	-4.031**
CONFIDENT	WARY	8.094***	1.480 NS	4.717***	2.414*
FRIENDLY	UNFRIENDLY	.696 NS	4.299**	13.038***	-6.994***
BRAVE	SCARED	4.965***	.803 NS	4.778***	1.739 NS
LAID-BACK	WORRIED	1.740 NS	-.049 NS	.488 NS	.326 NS
CALM	BOISTEROUS	-4.918***	.482 NS	-3.730**	-1.360 NS
AFFECTIONATE	COLD	.279 NS	5.101***	11.546***	-8.384***
STEADY	JUMPY	.558 NS	.608 NS	-.703 NS	-1.005 NS
NOSY	NOT BOTHERED	3.793**	2.547*	7.910***	-3.127*
STRONG	ANXIOUS	4.277**	-.840 NS	2.478*	2.391*
DOMINANT	SUBMISSIVE	5.721***	-.164 NS	5.657***	4.652***
ENERGETIC	LAZY	9.028***	-.942 NS	7.395***	2.308*
RELAXED	UPTIGHT	3.434**	.836 NS	1.348 NS	-1.036 NS
ALERT	UNALERT	3.650**	1.442 NS	3.808**	1.194 NS
GENTLE	HARD	-1.886 NS	2.342**	1.426 NS	-6.818***
COMFORTABLE	UNCOMFORTABLE	4.042**	-.502 NS	2.575**	-1.279 NS
QUIET	LOUD	-7.417***	2.076 NS	-2.668*	-1.821 NS
EASY GOING	STUBORN	.295 NS	1.779 NS	.760 NS	-3.947**
SLOW	FAST	-16.001***	4.908***	-4.948***	-3.103*

Appendix 25

Performance and degree of agreement

We correlated Kendall's W and mean performance values. There was a significant correlation between horses' performance and the observers' degree of agreement for stable 2, but not for stables 1 and 3.

Correlation of performance scores and MDS personality dimensions

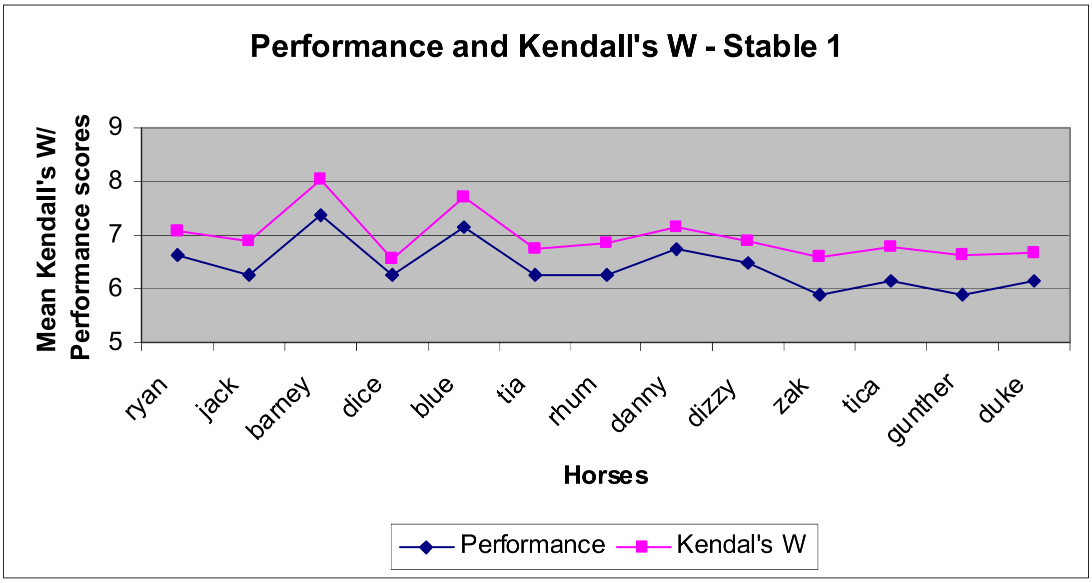
	Stable 1	Stable 2	Stable 3
Spearman's rho	.497 NS	.58*	.084 NS

*** P<0.001; **P<0.01; *P<0.05; NS – not significant

We also plotted mean performance scores and mean Kendall's W to see how the results are reflected in individual horses.

Stable 1

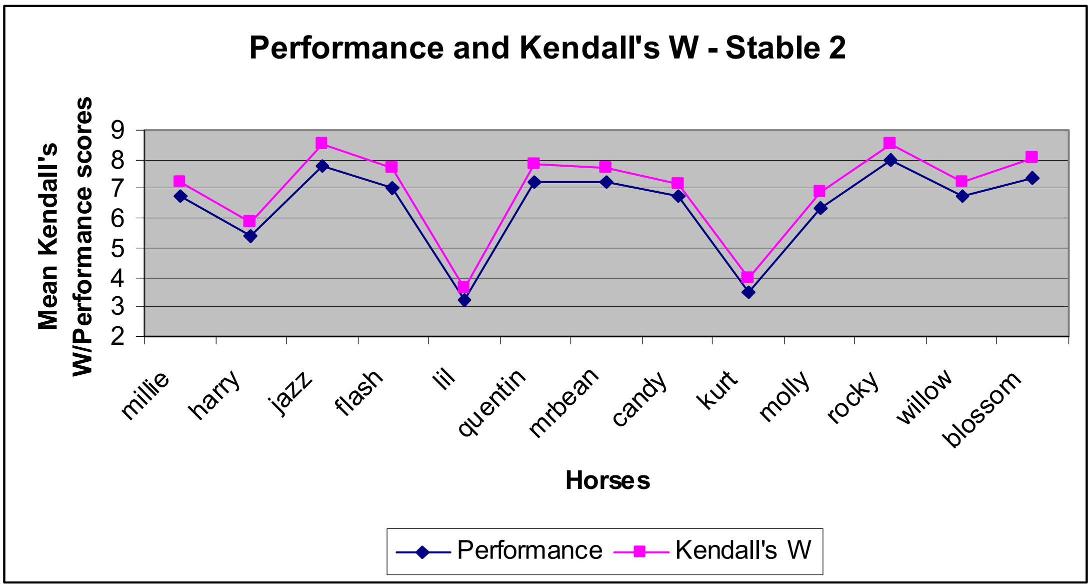
Even though the correlation is not significant, some horses e.g. Barney and Blue are high on performance scale and high on agreement, while Dice is low on both. Zak, Tica and Gunther do not show a relationship between the two sets of scores.



Performance and mean personality scores – Stable 1

Stable 2

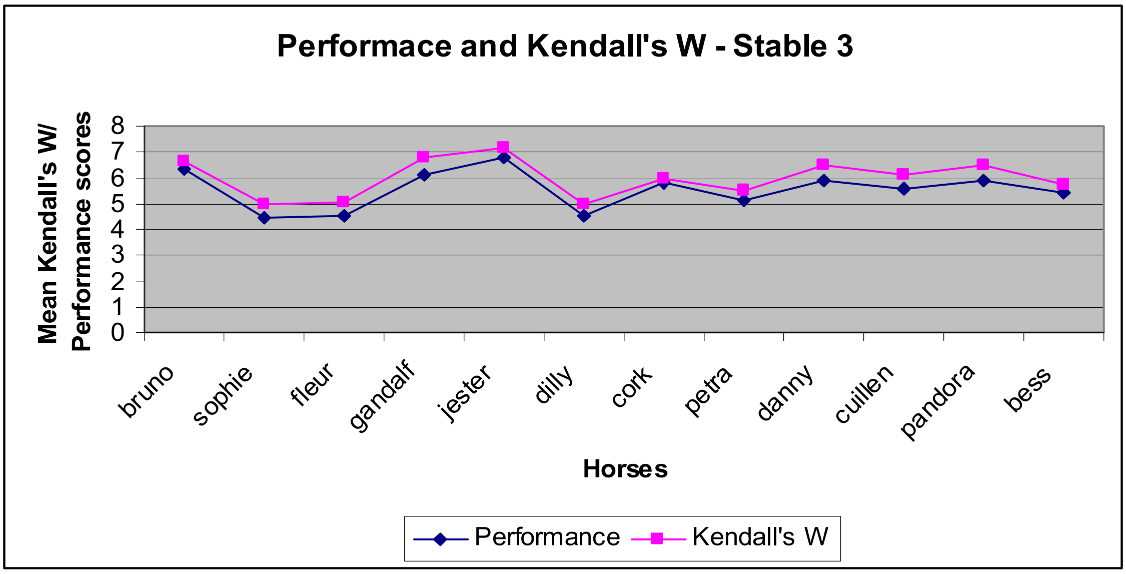
The correlation in Stable two is significant, which is illustrated by stacked line graph. Horses with high performance also have higher degree of agreement between the observers. For example, Lil has the lowest performance scores, and also the lowest degree of agreement between the observers. Rocky on the other hand has the highest performance scores and Kendall's W (Figure 8.31).



Performance and mean personality scores – Stable 2

Stable 3

Correlation for Stable 3 shows the most disharmonious relationship between the two lines. Even though the line graphs are stacked the discrepancy is manifested across most of the horses (Figure 8.32).



Performance and mean personality scores – Stable 3

Even though there is significant correlation between performance and the Kendall's W only in Stable 2, we can see some consistency between the two lines in Stable 1, but not for Stable 3.

Appendix 26

RATERS INFORMATION FORM – Study 1

Thank you for participating in the pilot study of Personality assessment in horses. Please fill in the questionnaire as correctly as possible. All the information will be kept strictly confidential and will be accessible only to the experimentor.

Name

Address

Phone number

Sex:
1. Male
2. Female

Date of birth

Occupation

Educational level:

1. University/ College degree
2. Other professional or technical qualification after leaving the school
3. Secondary school
4. Primary school
5. Other (please specify)

Are you a vegetarian?

1. Yes
2. No
3. Do not want to answer

If yes, which are the main reasons for not eating meat?

How long have you been working at the Riding School?

How long have you known the horses? Please name each horse and specify the period?

Appendix 27

ANIMAL EMPATHY SCALE (source: Paul, 2000)

Below is a list of statements that different people have made about the way they feel about animals, representing a range of different views.

Please indicate how strongly you agree or disagree with the following statements, by drawing a circle around the appropriate number on the agreement – disagreement scale. For example, if you think you agree with a statement fairly strongly, you might circle the 2 on the left-hand side of the scale:

Agree very strongly — Disagree very strongly

4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4

	Agree very strongly — Disagree very stron
1. So long as they are warm and well fed, I don't think zoo animals mind being kept in cages….	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
2. Often cats will meow and pester for food even when they are not really hungry………………	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
3. It upsets me to see animals being chased and killed by lions in wildlife programs on TV……..	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
4. The thought of calves being reared in veal crates really makes me feel sad……………………	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
5. Sad films about animals often leave me with a lump in my throat…………………………..	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
6. Animals deserve to be told off when they're not behaving properly………………………...	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
7. People are too concerned about the suffering of laboratory rats and mice……………………..	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
8. People who cuddle and kiss their pets in public annoy me……………………………………..	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
9. A friendly purring cat almost always cheers me up……………………………………………	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
10. It upsets me when I see helpless old animals…………………………………………………	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
11. Dogs sometimes whine and whimper for no real reason………………………………………	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
12. It makes me angry to think of the things that are done to laboratory animals………………...	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
13. I get very angry when I see animals being ill treated…………………………………………	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
14. It is silly to become too attached to one's pets………………………………………………...	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
15. Pets have a great influence on my moods……………………………………………………..	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
16. Sometimes I am amazed how upset people get when an old pet dies…………………………	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
17. Its silly to worry about how farm animals feel………………………………………………...	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
18. Seeing animals in pain upsets me……………………………………………………………...	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
19. People often make too much of the feelings and sensitivities of animals……………………..	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
20. I find it irritating when dogs try to greet me by jumping up and licking me………………….	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
21. I would always try to help if I saw a dog or puppy that seemed to lost……………………….	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
22. I hate to see birds in cages where there is no room for them to fly about……………………..	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
23. It upsets me to see farm animals in lorries going to slaughter………………………………...	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
24. I enjoy feeding scraps of food to the birds……………………………………………………	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
25. It makes me sad to see an animal on its own in a cage………………………………………..	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
26. I get annoyed by dogs that howl and bark when they are left alone…………………………..	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
27. I hate seeing pictures of animals used in scientific experiments………………………………	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4
28. Many people are over-affectionate towards their pets………………………………………...	4 – 3 – 2 – 1 – 0 – 1 – 2 – 3 – 4

Appendix 28

Observers' education – Study 1

80% of participants have either finished or were at University or College (Figure 9. 3).

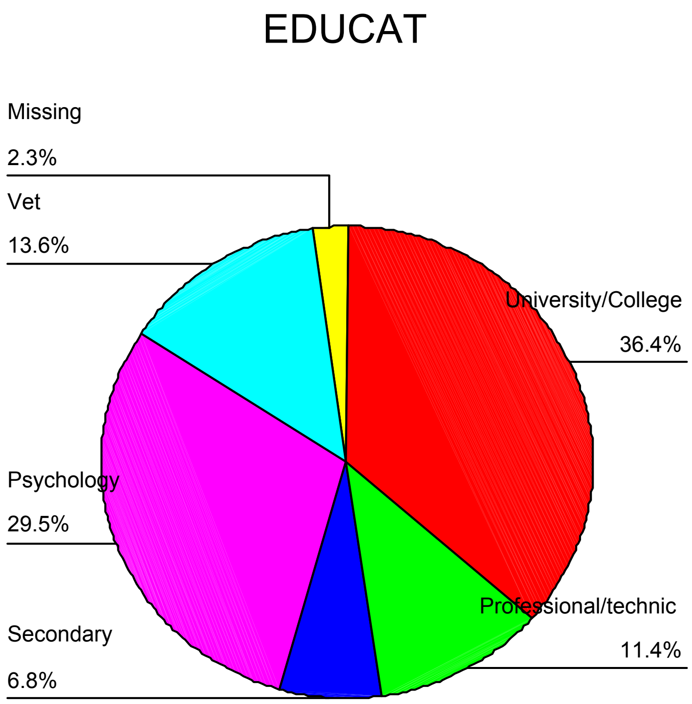


Figure 9.3
Level of education

Almost 30% were psychology students and almost 14% veterinary students.

Appendix 29

Kendall's W for unfamiliar observers high and low on the AES, Stable 1

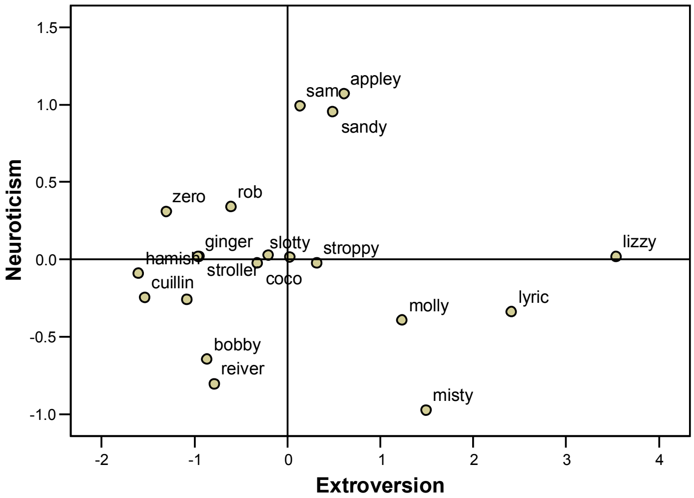
HORSE	Experienced W		Inexperienced W	
	High AES	Low AES	High AES	Low AES
1. HAMISH	.564***	.611***	.334***	.420***
2. STROLLER	.380***	.375***	.380***	.309***
3. CUILLIN	.641***	.608***	.479***	.513***
4. SAM	.397***	.320***	.417***	.318***
5. BOBBY	.518***	.548***	.465***	.593***
6. ZERO	.563***	.645***	.459***	.607***
7. FUDGE	.525***	.504***	.244 **	.497***
8. GINGER	.539***	.535***	.546***	.619***
9. SANDIE	.474***	.411***	.226 **	.419***
10. EDDIE	.459***	.494***	.173 NS	.350***
Mean Kendall's W	**0.506**	**0.5051**	**0.3723**	**0.4645**

Kendall's W for unfamiliar observers high and low on the AES, Stable 2

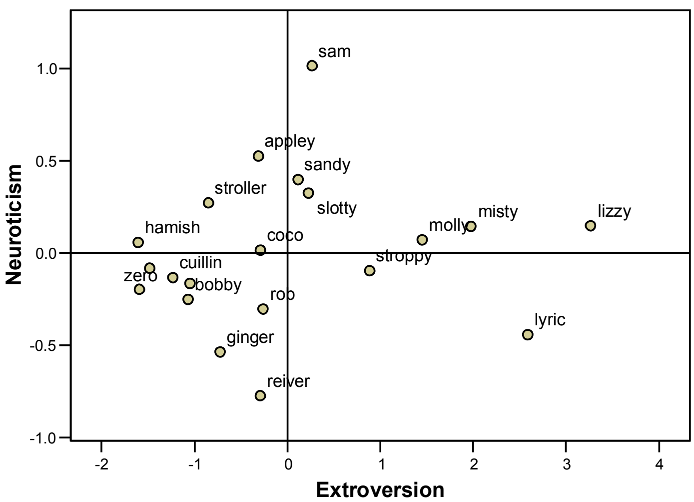
HORSE	Experienced W		Inexperienced W	
	High AES	Low AES	High AES	Low AES
1. LIZZY	0.684***	0.643***	0.464***	0.466***
2. APPLEY	0.539***	0.512***	0.624***	0.624***
3. MOLLY	0.303***	0.319***	0.317***	0.356***
4. COCO	0.355***	0.315***	0.431***	0.281***
5. LYRIC	0.585***	0.439***	0.563***	0.492***
6. STROPPY	0.243***	0.168*	0.407***	0.342***
7. REIVER	0.458***	0.242***	0.293***	0.299***
8. SLOTTY	0.277***	0.263***	0.435***	0.288***
9. MISTY	0.425***	0.486***	0.302***	0.538***
10. ROB	0.451***	0.268***	0.513***	0.502***
Mean Kendall's W	**0.432**	**0.3655**	**0.4349**	**0.4188**

Appendix 30

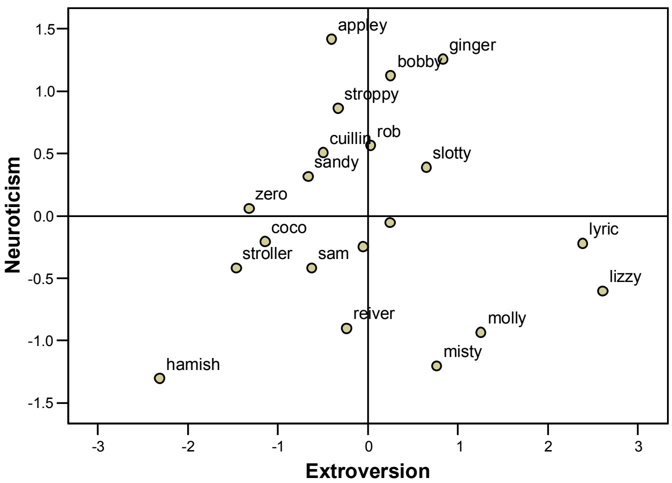
Experienced observers high on Animal empathy scale AES, Stables 1 and 2



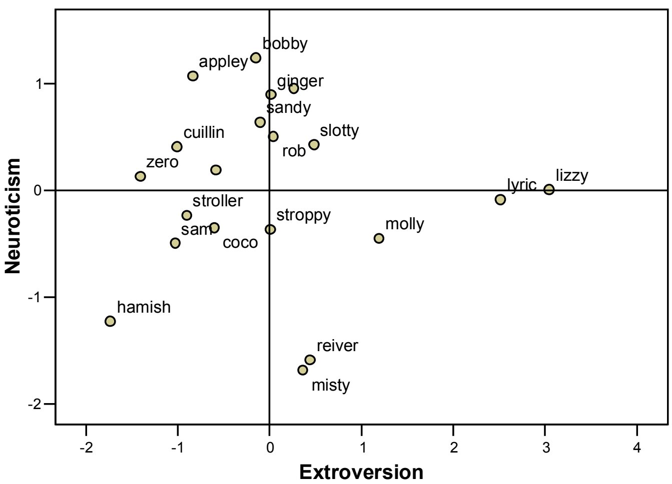
Experienced observers low on Animal empathy scale AES, Stables 1 and 2



Inexperienced observers low on Animal empathy scale AES, Stables 1 and 2

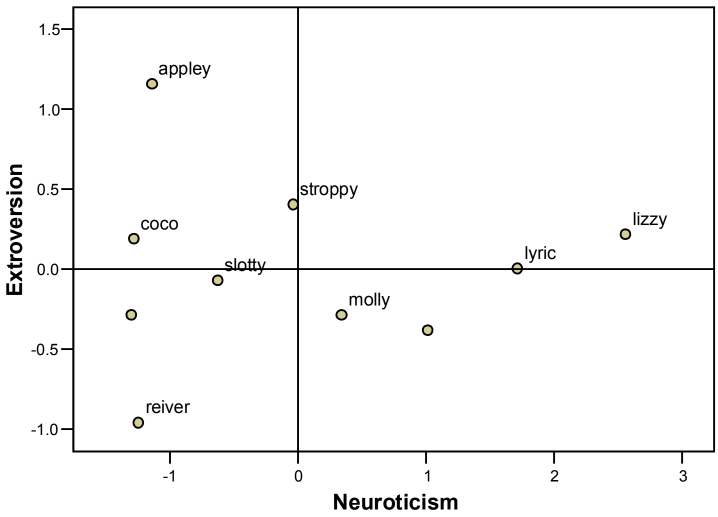


Inexperienced observers high on Animal empathy scale AES, Stables 1 and 2

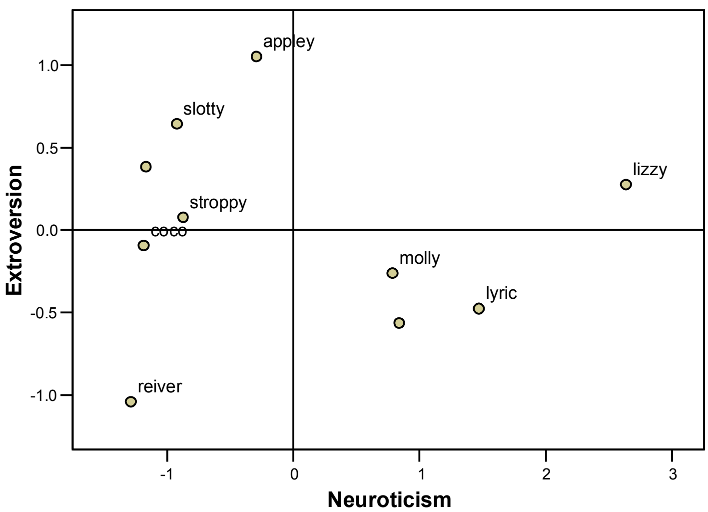


Appendix 31

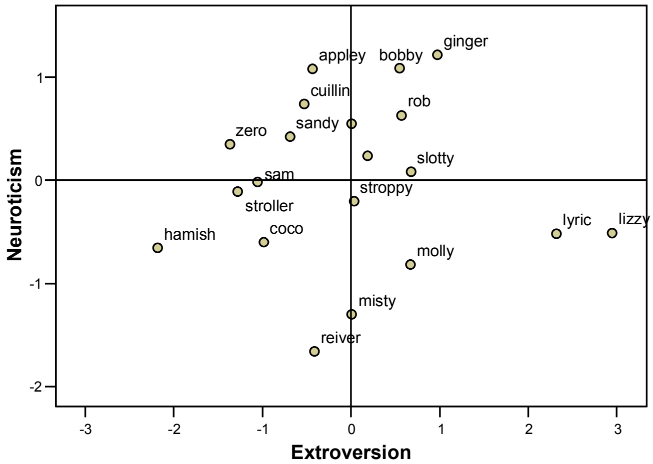
Experienced observers high on Neuroticism, Stable 2



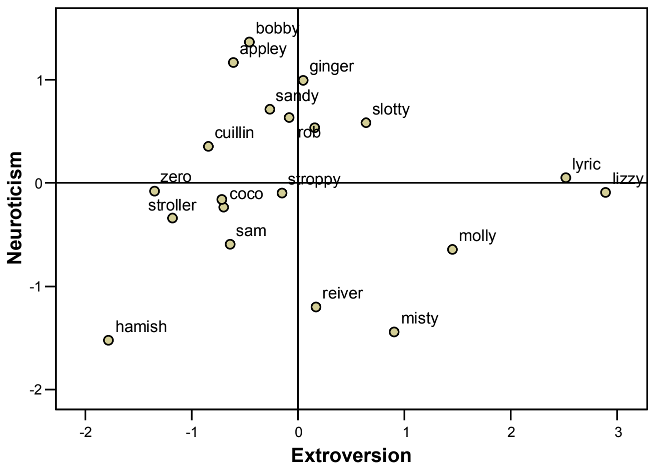
Experienced observers low on Neuroticism, Stable 2



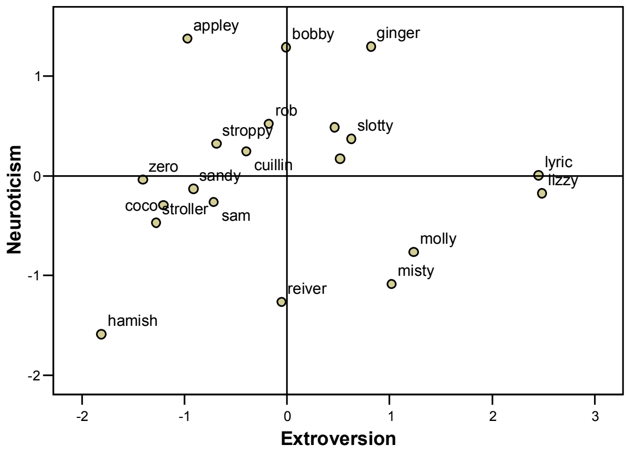
Inexperienced observers high on Extroversion, Stables 1 and 2



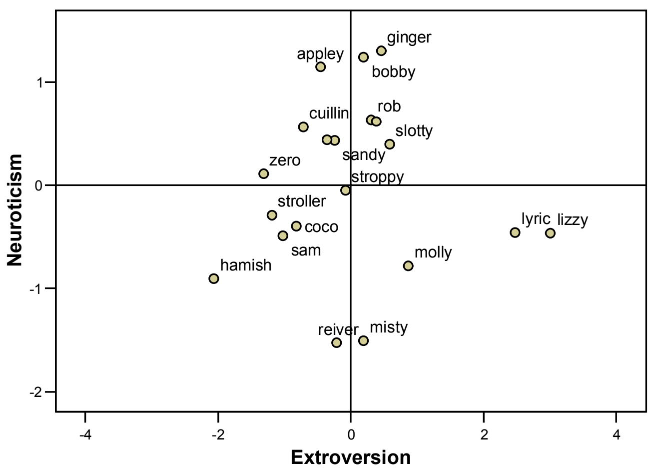
Inexperienced observers low on Extroversion, Stables 1 and 2



Inexperienced observers high on Neuroticism, Stables 1 and 2



Inexperienced observers low on Neuroticism, Stables 1 and 2



Appendix 32

Emotional Intelligence Questionnaire (from Schutte et al., 1998)

The following 33 statements provide you with an opportunity to describe yourself by indicating the degree to which each statement is true of the way you feel, think, or act most of the time and in most situations. Read each statement and decide which one of the five possible responses (***1. Strongly disagree; 2. Disagree; 3. Neither agree or disagree; 4. Agree; 5. Strongly agree***) best describes you. Mark your choices on the answer sheet by circling the number that corresponds to your answer. There are no "right" or "wrong" answers and no "good" or "bad" choices.

NAME: ______________________ **DATE:**

		Strongly Disagree				Strongly Agree
1.	I know when to speak about my personal problems to others.	1	2	3	4	5
2.	When I am faced with obstacles, I remember times when I faced similar obstacles and overcame them.	1	2	3	4	5
3.	I expect that I will do well on most things I try.	1	2	3	4	5
4.	Other people find it easy to confide in me.	1	2	3	4	5
5.	I find it hard to understand the non-verbal messages of other people.	1	2	3	4	5
6.	Some of the major events of my life have led me to re-evaluate what is important and not important.	1	2	3	4	5
7.	When my mood changes I see new possibilities.	1	2	3	4	5
8.	Emotions are one of the things that make my life worth living.	1	2	3	4	5
9.	I am aware of my emotions as I experience them.	1	2	3	4	5
10.	I expect good things to happen	1	2	3	4	5
11.	I like to share my emotions with others.	1	2	3	4	5
12.	When I experience a positive emotion, I know how to make it last.	1	2	3	4	5
13.	I arrange events others enjoy.	1	2	3	4	5
14.	I seek out activities that make me happy.	1	2	3	4	5
15.	I am aware of the non-verbal message that I send others.	1	2	3	4	5
16.	I present myself in a way that makes a good impression on others.	1	2	3	4	5
17.	When I am in a positive mood, solving problems is easy for me.	1	2	3	4	5
18.	By looking at their facial expressions, I recognise the emotions people are experiencing.	1	2	3	4	5
19.	I know why my emotions change.	1	2	3	4	5
20.	When I am in a positive mood, I am able to come up with new ideas.	1	2	3	4	5
21.	I have control over my emotions.	1	2	3	4	5
22.	I easily recognize my emotions as I experience them.	1	2	3	4	5
23.	I motivate myself by imagining a good outcome to tasks I take on.	1	2	3	4	5
24.	I compliment others when they have done something well.	1	2	3	4	5
25.	I am aware of the non-verbal messages other people send.	1	2	3	4	5
26.	When another person tells me about an important event in his or her life, I almost feel as though I have experienced the event myself.	1	2	3	4	5
27.	When I feel a change in emotions, I tend to come up with new ideas.	1	2	3	4	5
28.	When I am faced with a challenge, I give up because I believe I will fail.	1	2	3	4	5

29.	I know what other people are feeling just by looking at them.	1	2	3	4	5
30.	I help other people feel better when they are down.	1	2	3	4	5
31.	I use good moods to help myself keep trying in the face of obstacles.	1	2	3	4	5
32.	I can tell how someone is feeling from their tone of voice.	1	2	3	4	5
33.	It is difficult for me to understand why people feel the way they do.	1	2	3	4	5

Appendix 33

Kendall's W and Chi-square for familiar and unfamiliar observers, low and high on EI scale, Study 2, Stable 1

HORSE	Familiar W (γ^2)		Unfamiliar W (γ^2)	
	Low EI	High EI	Low EI	High EI
1. RYAN	.622** 63.42	.447 ** 60.853	.472*** 63.761	.551*** 74.359
2. JACK	.710*** 72.406	.656*** 89.268	.373 ** 50.337	.477*** 64.350
3. BARNEY	.719*** 73.356	.647*** 87.927	.241 NS 32.477	.551*** 74.359
4. DICE	.621 ** 63.306	.333 NS 45.248	.246 NS 33.161	.425*** 57.338
5. BLUE	.603 ** 61.495	.597*** 81.184	.328** 44.287	.391** 52.783
6. TIA	.467 NS 47.588	.596*** 81.004	.448** 48.396	.437*** 59.041
7. RHUM	.787*** 80.312	. .548*** 74.478	.390** 52.606	.366** 49.456
8. DANNY	.722*** 73.650	.483*** 65.735	.286 NS 38.548	.391** 52.768
9. DIZZY	.612 ** 62.431	.428** 58.144	.308* 41.559	.229 NS 30.967
10. ZAK	.659*** 67.219	.645*** 87.759	.552*** 74.502	.533*** 71.901
11. TICA	.773*** 78.875	.682*** 92.795	.388** 52.363	.549*** 74.067
12. GUNTHER	.808*** 82.453	.777*** 105.622	.377** 50.930	.368** 49.630
13. DUKE	.680*** 69.336	.647*** 87.988	.385 ** 52.019	.452*** 60.955
Mean Kendall's W	**0.675**	**0.578**	**0.368**	**0.44**
Chi-square mean	**68.911**	**78.308**	**48.842**	**59.382**

Kendall's W and Chi-square for familiar and unfamiliar observers, low and high on EI scale, Study 2, Stable 2

HORSE	Familiar W (γ^2)		Unfamiliar W (γ^2)	
	Low EI	High EI	Low EI	High EI
1. MILLIE	.508* 51.771	.572*** 97.186	.478*** 64.566	.427*** 57.581
2. HARRY	.639*** 65.187	.472*** 80.282	.613*** 82.750	.307* 41.425
3. JAZZ	.871*** 88.799	.765*** 130.068	.610*** 82.329	.238 NS 32.163
4. FLASH	.816*** 83.236	.733*** 124.553	.447*** 60.405	.304* 41.020
5. LIL	.681*** 69.438	.382*** 64.928	.231 NS 31.156	.310* 41.820

6. QUENTIN	.819*** 83.522	.566*** 96.262	.457*** 61.688	.555*** 74.957
7. MR. BEAN	.829 ** 56.403	.448*** 76.230	.417*** 56.352	.397** 53.600
8. CANDY	.594 ** 60.623	.373*** 63.354	.274 NS 37.040	.441*** 59.599
9. KURT	.704*** 71.807	.420*** 71.438	.759*** 102.484	.472*** 63.730
10. MOLLY	.777*** 79.287	.474*** 80.587	.632*** 85.309	.503*** 67.855
11. ROCKY	.688*** 70.178	.472*** 80.215	.341* 45.993	.269 NS 36.286
12. WILLOW	.712*** 72.579	.468*** 79.476	.326* 44.058	.284 NS 38.318
13. BLOSSOM	.846*** 86.270	.620*** 105.347	.392** 52.953	.224 NS 30.264
Mean Kendall's W	0.729538	0.520385	0.459769	0.363923
Chi-square mean	72.23846	88.45585	62.08331	49.12446

Kendall's W and Chi-square for familiar observers, low and high on EI scale, Study 2, Stable 3

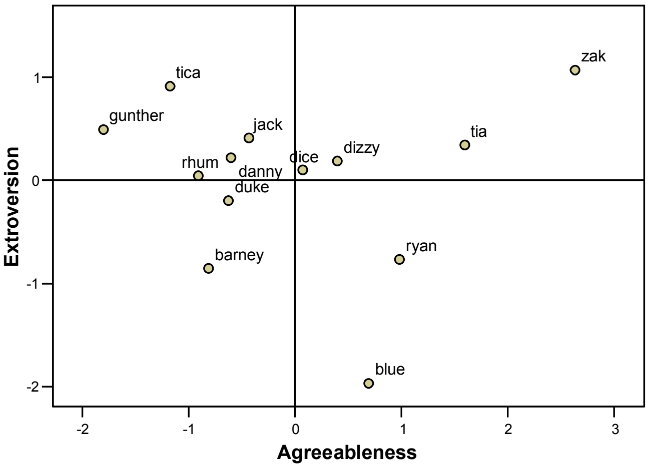
HORSE	**Familiar W (γ^2)**	
	Low EI	High EI
1. BRUNO	.474 NS 48.342	.436*** 89.002
2. SOPHIE	.617** 62.896	.648*** 132.291
3. FLEUR	.698*** 71.146	.549*** 111.927
4. GANDALF	.670*** 68.366	.760*** 155.039
5. JESTER	.421 NS 42.965	.538*** 109.735
6. DILLY	.529* 53.971	.476*** 97.183
7. CORK	.491* 50.128	.248* 50.522
8. PETRA	.683*** 69.696	.419*** 71.275
9. DANNY	.686*** 69.991	.661*** 134.844
10. CUILLEN	.619** 63.173	.663*** 135.337
11. PANDORA	.689*** 70.295	.659*** 134.479
12. BESS	.471 NS 48.050	.391*** 79.743
Mean Kendall's W	0.587333	0.563636
Chi-square mean	59.91825	108.4481

Appendix 34

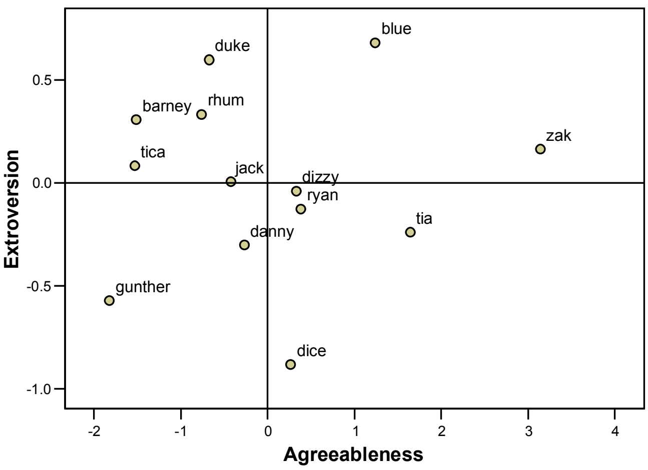
MDS Familiar observers Low and High on EI

Stable 1

Low EI

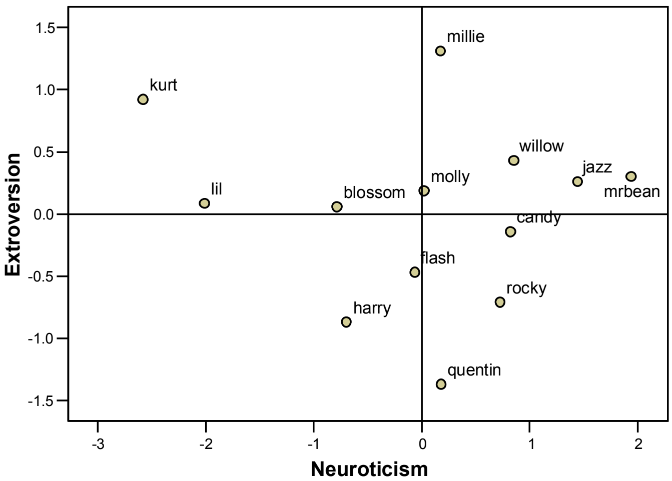


High EI

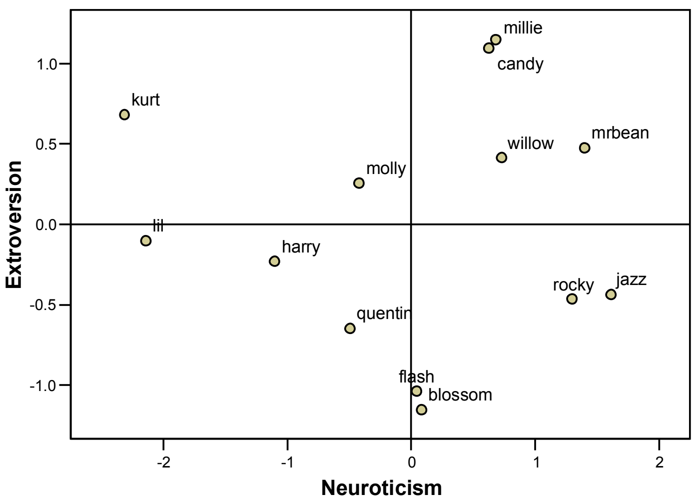


Stable 2

Low EI

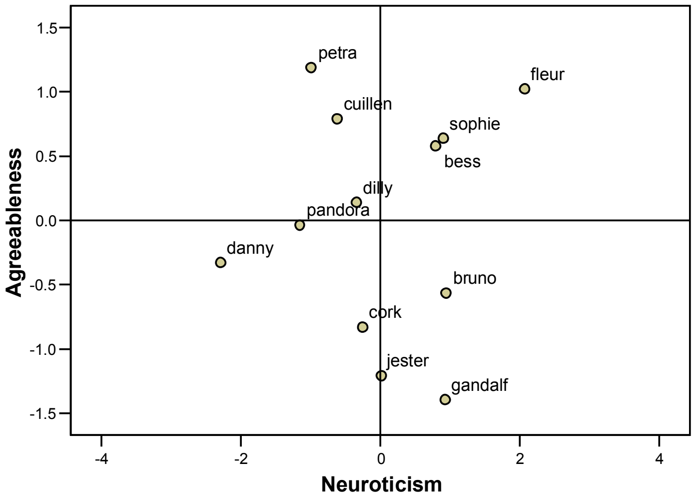


High EI

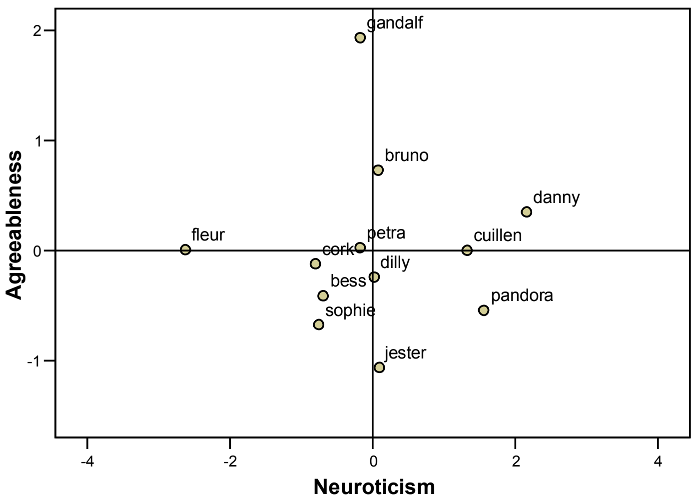


Stable 3

Low EI



High EI



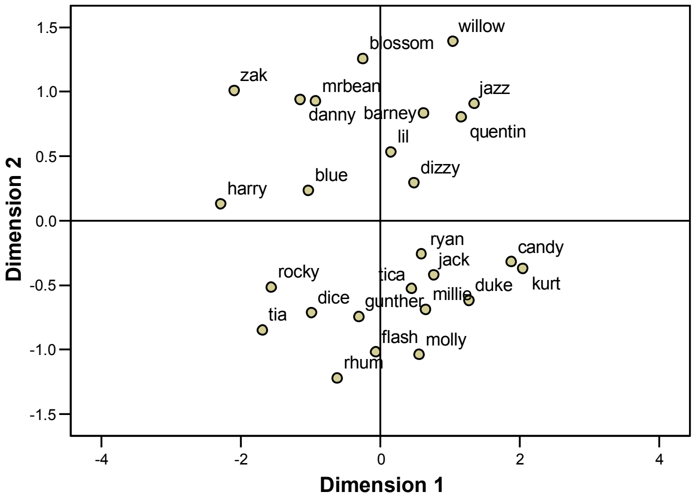
Appendix 35

Unfamiliar Low and High on EI – All stables

Low EI

Derived Stimulus Configuration

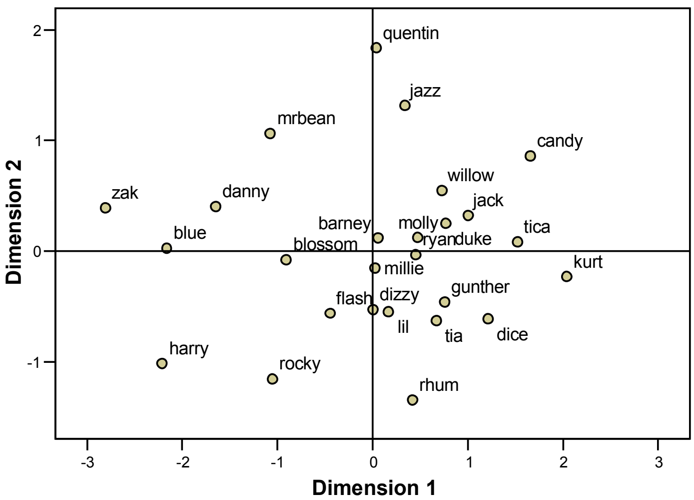
Euclidean distance model



High EI

Derived Stimulus Configuration

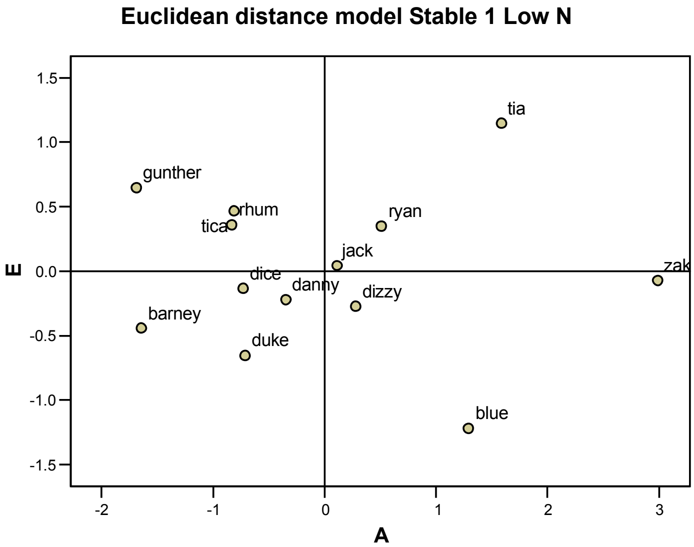
Euclidean distance model

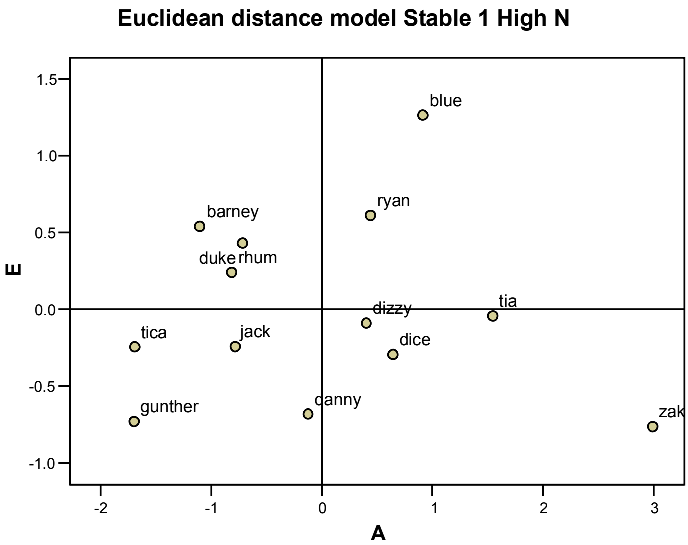


Appendix 36

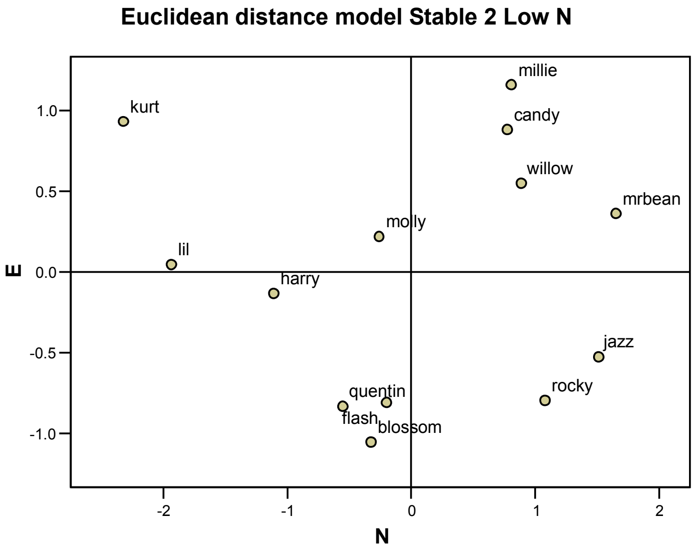
Observers low and high on Neuroticism dimension

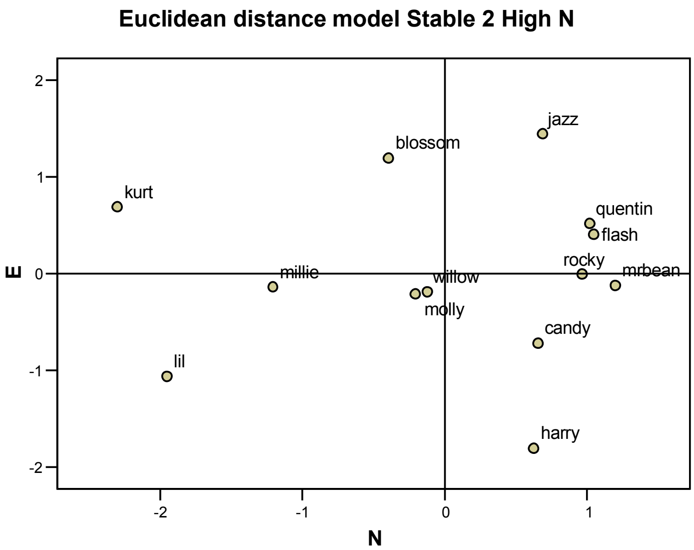
Familiar observers
Stable 1



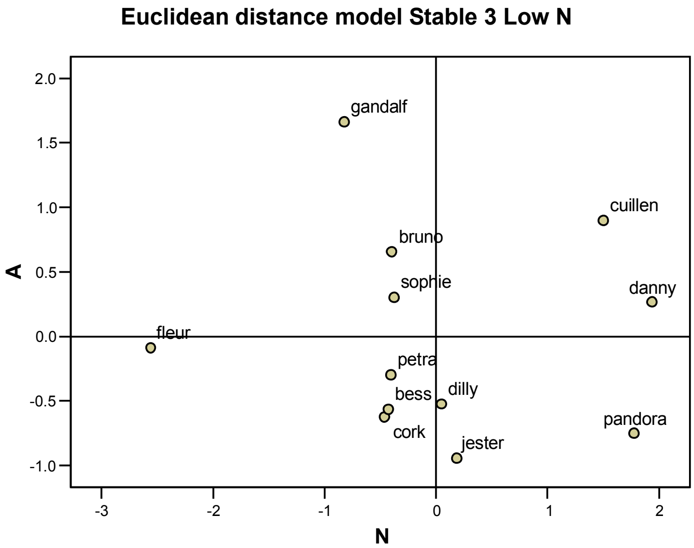


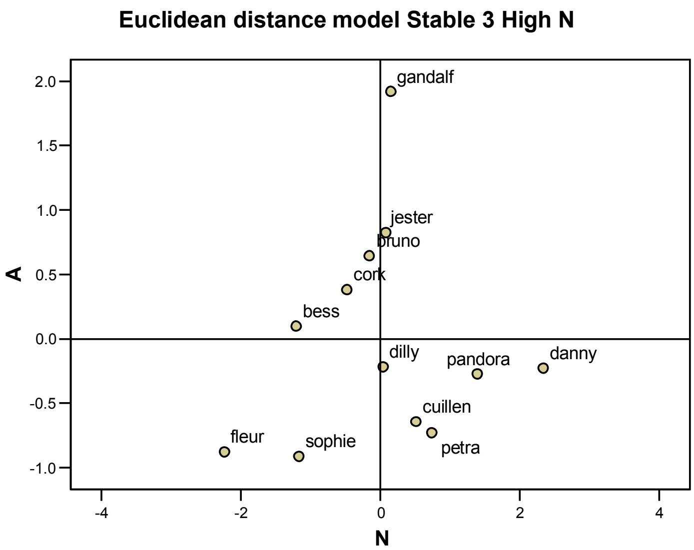
Stable 2


Euclidean distance model Stable 2 Low N
millie
kurt
candy
willow
mrbean
molly
lil
harry
jazz
quentin
rocky
flash
blossom
1.0
0.5
0.0
-0.5
-1.0
E
-2
-1
0
1
2
N


Euclidean distance model Stable 2 High N
jazz
blossom
kurt
quentin
flash
rocky
millie
willow
mrbean
molly
candy
lil
harry
2
1
0
-1
-2
E
-2
-1
0
1
N

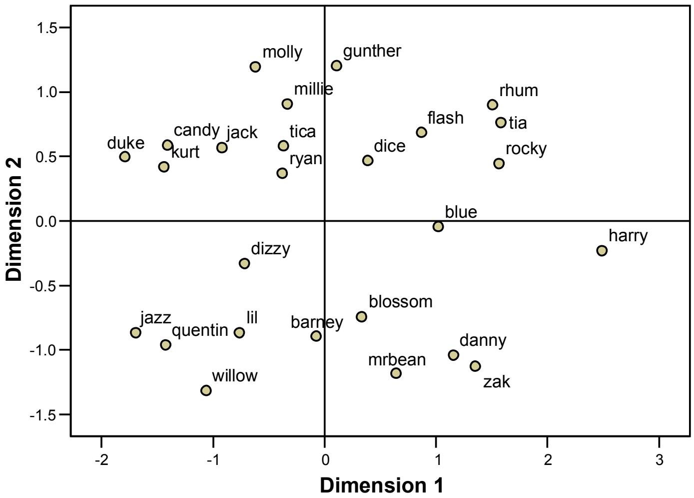
Stable 3


Euclidean distance model Stable 3 Low N
gandalf
cuillen
bruno
sophie
danny
fleur
petra
bess
dilly
pandora
cork
jester
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
A
-3
-2
-1
0
1
2
N

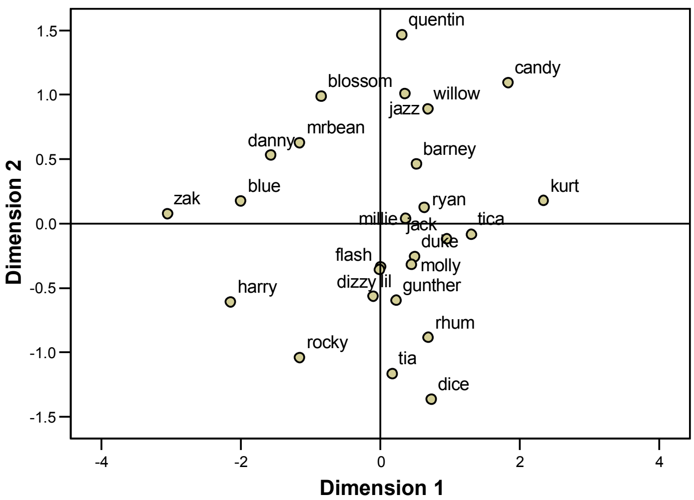

Euclidean distance model Stable 3 High N
gandalf
jester
bruno
cork
bess
dilly
pandora
danny
cuillen
petra
fleur
sophie
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
A
-4
-2
0
2
4
N

Unfamiliar observers – Stables 1 and 2

Euclidean distance model Unfamiliar observers Low N



Euclidean distance model Unfamiliar observers High N

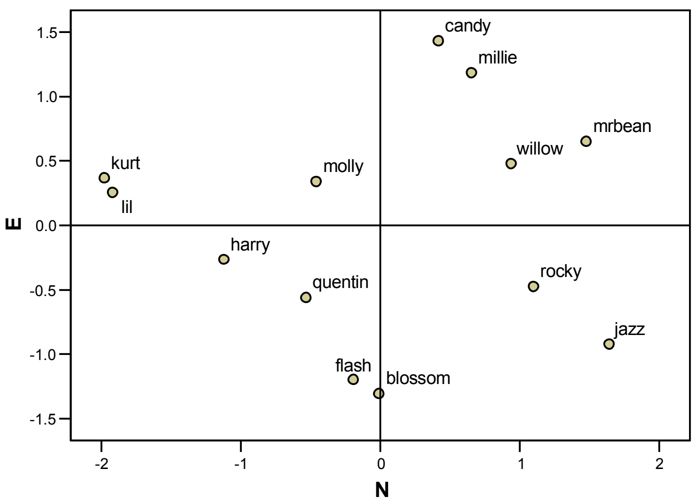


Appendix 37

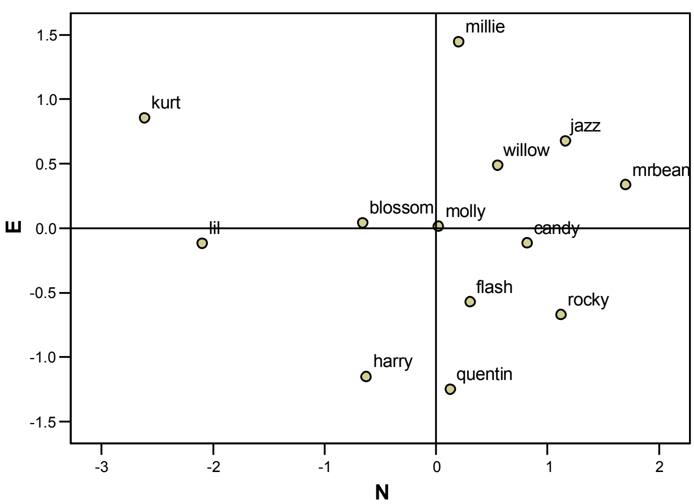
Observers low and high on Extroversion dimension

Familiar observers

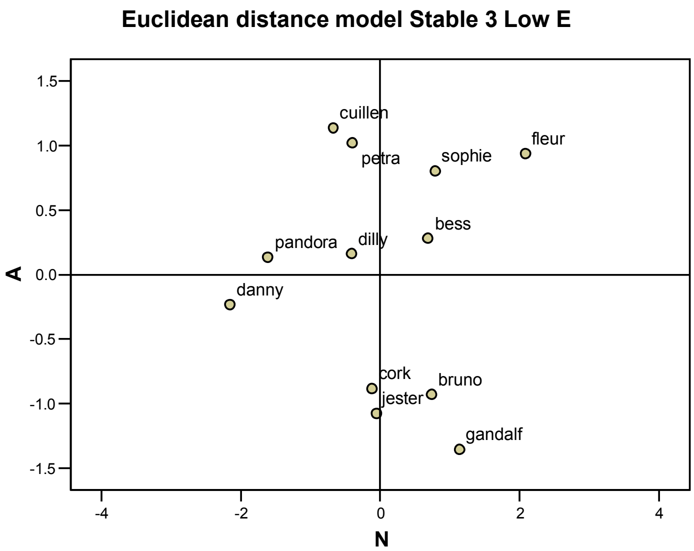
Stable 2

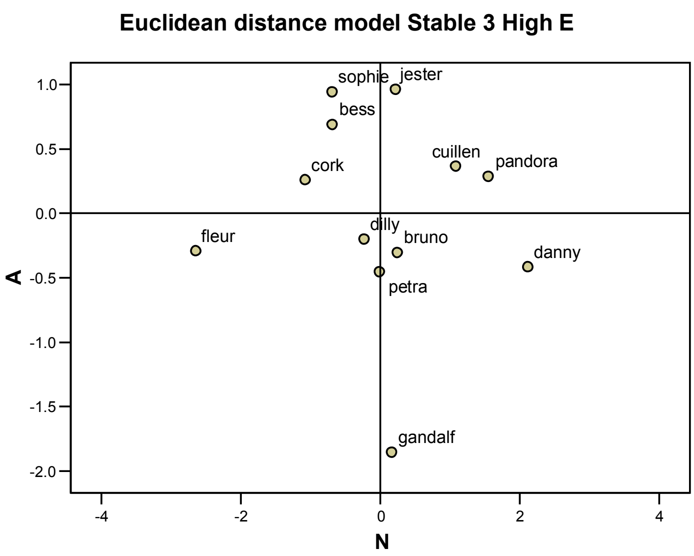


Euclidean distance model Stable 2 High E



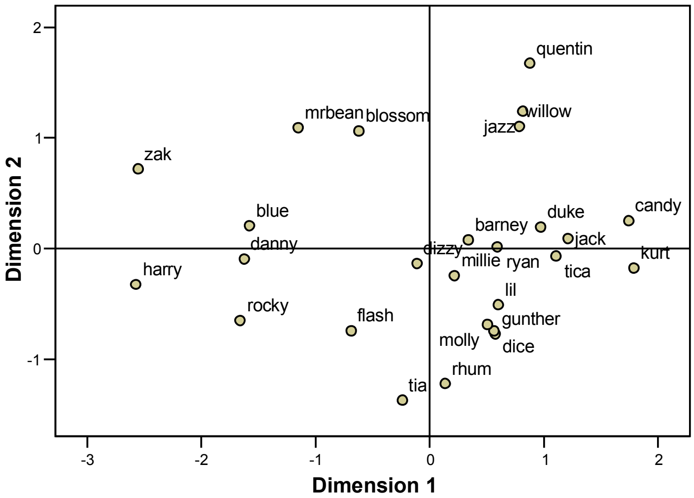
Stable 3


Euclidean distance model Stable 3 Low E
cuillen
petra
sophie
fleur
bess
pandora
dilly
danny
cork
bruno
jester
gandalf
A
N
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-4
-2
0
2
4


Euclidean distance model Stable 3 High E
sophie
jester
bess
cuillen
pandora
cork
fleur
dilly
bruno
danny
petra
gandalf
A
N
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-4
-2
0
2
4

Unfamiliar observers

Euclidean distance model Unfamiliar observers Low E



Euclidean distance model Unfamiliar observers High E