Australian Government
Department of Agriculture

AUSTRALIAN Pork™

Identify and Integrate Measures of Animal Welfare that Meet the Needs of Animals and Society – Part 1

Final Report
APL Project 2011/1036.421

February 2013

CSIRO Animal, Food and Health Sciences
Drewe Ferguson & Dr Ian Colditz
671 Sneydes Road
Werribee Vic 3030
Murdoch University
Teresa Collins
90 South Street
Murdoch WA 6150
Lindsay Matthews & Associates Research International
Dr Lindsay Matthews
The University of Melbourne
Animal Welfare Science Centre
Prof Paul Hemsworth
Faculty of Land & Food Resources
Vic 3010

Acknowledgements

This project is supported by funding from Australian Pork Limited and the Department of Agriculture.

Executive Summary – Part 1

As with the treatment of domesticate animals in other settings, there is increasing societal concern about the treatment of livestock on farms and their quality of life within production environments. Viable livestock farming requires practices that are not only productive, profitable and sustainable but that also fit with society's expectations on ethical dimensions such as animal welfare. Transparent demonstration of how these expectations have been met will be paramount in the future.

To establish whether an animal's physical and emotional needs are being met requires a detailed assessment of its welfare. Welfare assessment is a major challenge as the utility of any assessment methodology depends on the specific situation under examination and the ethical views held by the stakeholder group seeking the assessment. The purpose of this review is to explore the subject of welfare assessment further by examining the conceptual frameworks, complexities and methodologies applied to the task.

The specific aims of Part 1 of the review are to:

1. undertake a comprehensive analysis of the scientific literature on welfare measures and assessment methods to identify the most credible scientific measures that could be developed into a uniform field index and,
2. recommend where further research to validate welfare concepts and methodologies is required.

The review describes the historical and current contexts of animal welfare and the commonly used conceptual frameworks for its assessment. The broad categories of measures used in science and in welfare assessment and assurance systems used on-farm are then reviewed.

The key conclusions are:

- The pre-eminent concepts of good welfare employed today encompass biological functioning, affective states and naturalness.
- There is ongoing need to demonstrate the validity of welfare measures (i.e. show that measures accurately reflect an animal's welfare state or the definition of a good animal welfare state held by the stakeholder group(s) seeking the assessment .
- Welfare assessment is an evaluative process in which values influence the choice of the conceptual framework and consequently the measures, their interpretation and their weighting when the measures are combined in any legislative standard, QA assessment system, welfare management tool or research methodology.
- A better understanding of the emotional range and valence in livestock species is required, as is a better understanding of the consequences of injury and illness for animal emotions. To that end, the ongoing development and validation of behavioural and cognitive methodologies is essential. This goal could be greatly enhanced through the application of neuroscience disciplines to studies of livestock behaviour and emotions.
- There is no one comprehensive, fully-validated system for on-farm welfare assessment that accommodates the diversity of species, production environments and animal management

systems practiced in Australia. However, research has shown that assessments based on combinations of health and production data together with observation of behaviour and physical appearance of animals within a group offer reliable and feasible tools for the assessment of welfare. The strategic combination of input or resource-based and outcome or animal-based measures will also be important, particularly in the context of welfare risk assessment and risk management (e.g. assessments of pasture/forage availability and body condition score in cattle and sheep).

- Efforts should be directed at improving the practicability of welfare assessment systems within the various livestock enterprises. Specifically, further effort is needed to find ways to improve the reliability whilst reducing the complexity and invasiveness of methodologies. The development and application of remote automated data capture systems could be valuable in both extensive and intensive animal production systems.

In conclusion, the development of appropriate welfare assessment methodologies that are credible to all stakeholders will be built on a better understanding of:

- changes in biological functioning and fitness including changes in physical health that correspond with different levels of welfare;
- the capacities of livestock to experience negative and positive mental states and associated levels of welfare;
- the ways that separate measures and welfare attributes can be weighted and integrated to give an overall index of welfare; and
- how these can be practically and reliably implemented in the production environment.

Part 2 of the review describes in detail current welfare assessment schemes and looks at the challenges faced in devising a comprehensive assessment program suitable for cross-sectional application. It proposes a new unified field index for implementation through a process of risk assessment, risk management and benchmarking to provide a welfare management and assessment tool for use across Australia's livestock industries.

1. Scope of the Review

Increasing societal interest in the treatment of animals used in livestock production systems has led to recognition of the need for ongoing improvements in animal welfare. Concern about an animal's quality of life is placing an increasing onus on those who husband animals to provide for the animal's physical and emotional needs, and to provide information about husbandry practices to the public. To describe how adequately the needs of animals are met requires a detailed assessment of the animals in their environment. This is a major challenge on a number of levels as the utility of any welfare assessment methodology will depend on the specific situation being assessed and the ethical views held by the stakeholder group seeking the assessment. The purpose of this review is to explore the subject of welfare assessment by examining the conceptual frameworks, complexities and methodologies that are applied to the task of assessing welfare in farm animals.

The review will be presented in two parts. In the Part 1, the specific aims are to:
(1) undertake a comprehensive review of the scientific literature on welfare measures and assessment methodologies in order to identify the most credible scientific measures that could be developed into a uniform field index and
(2) recommend where further research to validate welfare concepts and methodologies is required.
Part 2 of the review examines novel methods for integrating these measures into a welfare index.

The review examines:

- societal concerns about animal welfare
- conceptual frameworks for assessing welfare
- measures of welfare, and
- current methods for assessing welfare in research and on farm

The second part of the review addresses approaches to integrating measures into a unified field index of farm animal welfare and proposes a unified filed index for on=farm application across Australia's livestock sectors.

2. Setting the Scene

Awareness and concern about animal welfare usually arise when the actions of humans intersect with the lives of animals, especially those animals born into our care and responsibility. When we approach the issue of animal welfare we do so from a number of perspectives that are influenced by evidence, by values and by attitudes about how animals ought to be treated. For each of us, these viewpoints can be considered to be personal anchoring points or landmarks from which we attempt to triangulate and fix on animal welfare, just as a GPS unit uses satellites to fix a position on the ground.

If you have ever driven down a country road watching a rainbow dance across the landscape, then stopped to take a photo, you will have witnessed a surprising phenomenon. Without the motion of the car, the rainbow becomes dull and shallow, and from some vantage points is not visible at all. So it is with animal welfare that as we move between various values, attitudes and evidence, the depth, vividness and character of the issue changes and are envisioned in greater detail and complexity. Projections from the landmarks of evidence, values and attitudes create the construct we call animal welfare that, like a hologram, appears to differ from each new perspective. As we move between landmarks we can see suffering, happiness, utility, fitness, pain, health, disease, purpose, arousal, depression, anxiety, exploitation and so on in many depths and intensities. From a single viewpoint, we sometimes understand very little at all about animal welfare and the acceptability of animal use practices.

It follows then, that each 'sighting' of animal welfare, such as say the perception of suffering, can be disaggregated into its constitutive elements of evidence, values and attitudes. Conversely, a single piece of evidence, say a cortisol response, can project to multiple aspects of the animal welfare hologram. Thus when interpreted through differing ideas of biological function, of affective states, or of naturalness, the datum may map to differing places within the animal welfare construct. Without knowledge of how the perception of welfare deconstructs into its constitutive elements, we are ignorant of how the perception arose in the first place and of what it tells us about the whole of which it is a part.

The role of science in the assessment of animal welfare has been discussed in detail in the literature (Barnett and Hemsworth, 2009; Croney et al., 2012; Sandoe et al., 2004; Tannenbaum, 1991). While scientific methods provide an objective way of collecting evidence, it is well recognized that the interpretation of data is influenced by values, as indeed is the initial choice of measures used to make the welfare assessment. The fact that scientific processes are underpinned with values does not diminish their contribution to welfare assessment but, as noted by the above authors, highlights the importance for scientists, like other stakeholders involved in animal welfare assessment, to clearly enunciate the value framework underpinning the interpretation of evidence.

Perceptions, understanding and interpretations of animal welfare are influenced by:

- Evidence
- Values, and
- Attitudes to animals

Scientific approaches to animal welfare are also influenced by these factors

2.1 Brief Historical Overview

The most prominent development and indeed the primary foundation of animal welfare is the fundamental acceptance that animals are sentient beings that feel emotions and are aware of their world through subjective experience. While this may seem to be a modern concept, this is not the case, as there was recognition of animal sentience in the writings from the Renaissance period spanning the 14-17th centuries (Duncan, 2006). Unfortunately, a consideration of animal feelings was not always reflected in the treatment of animals during this period. Significant momentum occurred during the 18th century when influential philosophers and social reformers challenged contemporary attitudes to animals. Preeminent amongst these reformers was Jeremy Bentham who asserted that "The question is not, Can they reason? Nor, Can they talk? But, Can they suffer?" Bentham's seminal question suggested that the capacity to suffer might be a sufficient criterion to entitle animals to legal rights. In the present day, this basic question continues to lie at the heart of animal welfare science where our challenge is to determine how best to quantify emotional states of animals and to ascertain the significance of these welfare states in relation to our duty of care in animal management. The last three decades have seen a significant expansion in efforts to understand the emotional repertoire of farm animals (Duncan, 2006). Most emphasis has been on quantifying negative or unpleasant emotional states (e.g. pain, fear, hunger). However, with an increasing community interest towards not merely minimising suffering in domestic animals, but also enhancing pleasure in these animals (Tannenbaum, 2001), there is now increasing attention on examining positive affective states and their importance within the lives of farm animals. We anticipate that an improved capacity to assess affective states of farm animals will change our biological understanding of the scope of animal experiences and have a significant bearing on future animal welfare policy and also lead to the development or refinement of some new or alternative animal production systems.

In terms of historical impact, two significant events occurred during the mid nineteen sixties. The first of these was the publication of *Animal Machines* by Ruth Harrison in 1964 which challenged the intensification of animal production and the concept of "factory farming". Secondly, in response to the marked increase in public concerns about intensive animal farming that followed publication of the book, the UK government established the Brambell Committee in 1965. The terms of reference of this committee were relatively simple; "To examine the conditions in which livestock are kept under systems of intensive husbandry and to advise whether standards ought to be set in the interests of their welfare, and if so what should they be?" (Brambell Committee, 1965). The full effects of the committee's report and their recommendations on UK and European animal welfare policy and legislation were both profound and enduring (Veissier et al., 2008) with consequences in other industrialized countries, including Australia. Furthermore, the report also strongly influenced the subsequent direction of animal welfare science at the time (Keeling et al., 2011).

Recognition and indeed advocacy of animal sentience was prominent in the views of the Brambell Committee which stated that animal welfare should encompass both the physical and emotional well being of the animal. They went on to say; "Any attempt to evaluate welfare therefore must take into account the scientific evidence concerning the feelings of animals that can be derived from their structure and functions and also from their behaviour" (Brambell Committee, 1965). Whilst consensus about the definition of animal welfare remains somewhat elusive, there are signs of some convergence (refer Section 3). Notwithstanding this, one could argue that the Committee's definition was certainly prescient.

- Recognising the sentience of animals is central to appreciating that they can suffer
- The capacity of animals to suffer has been recognised since the renaissance period
- A moral responsibility on humans to minimize suffering in animals has been acknowledged since at least the later 1700s
- There is increasing community interest towards not merely minimising suffering in domestic animals, but also enhancing pleasure in these animals

2. Conceptual Frameworks for Animal Welfare

A lesson from the historical changes in concepts of animal welfare is that current perceptions of what constitutes animal welfare are likely to also change. A common strategy to address the dilemma of how to assess animal welfare is to start by strict definition of what animal welfare is. This strategy has utility in that it helps sets the boundaries for what needs to be considered. A contemporary example is the definition adopted by European Food Safety Authority Panel on Animal Health and Welfare.

"Welfare refers to the state of an individual as it attempts to cope with its environment. Effects on welfare include changes in health, mental functioning, positive and negative feelings, physiological and behavioural responses and injuries."

The definition encompasses the range of states of the animal generally considered to be associated with its welfare. An obvious shortcoming of this type of definition is the absence of an explicit framework for interpreting evidence that falls within the parameters articulated in the definition. Three interpretive frameworks are commonly adopted and, even when used in combination tend to be weighted differently depending on the perspective of the assessor, the assessment methodology or the purpose for which the assessment is being undertaken. These three frameworks are:

- Biological functioning – normality as evidenced through measures of behaviour, physiology, health and productivity
- Affective states – as evidenced through measures of abnormal behaviours, affective states (positive and negative feelings) and cognitive function, and
- Naturalness – as evidence by attributes of the animal, or telos, in particular normal behavioural repertoires, and by attributes of its environment, in particular congruence between the extant production environment and a sometimes notional, preconceived ideal environment for the animal's species

Schematically the 3 domains are typically depicted in a Venn diagram as overlapping circles in which good welfare occurs within the common area as shown in Figure 1. Some alignment of disciplinary expertise and personal experience with each of the interpretive frameworks can also be recognised, with farmers and veterinarians often emphasizing biological functioning, cognitive and behavioural scientists often emphasizing mental functioning, and philosophers and animal rights proponents often emphasizing naturalness and integrity. The importance of insights drawn from each perspective to the appraisal of animal welfare and to ethical judgements is also well recognised. The interpretive frameworks are not necessarily competing views vying for dominance but complementary understandings that can potentially contribute to a more sophisticated and nuanced appreciation of animal welfare.

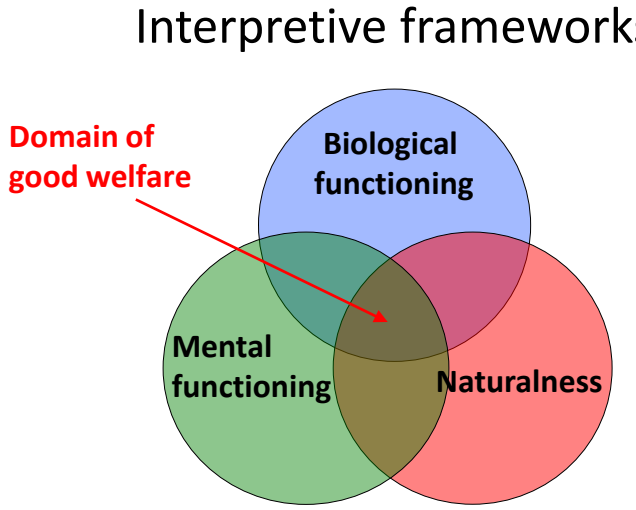


Figure 1: Schematic representation of three common conceptual frameworks for interpreting the welfare state of animals

The so-called 'Five Freedoms', that is:

1. Freedom from hunger and thirst,
2. Freedom from discomfort,
3. Freedom from pain, injury and disease,
4. Freedom to express normal behaviour, and
5. Freedom from fear and distress (FAWC, 1993)

They are included to varying degrees in each of three animal welfare concepts. While most would accept that these freedoms are necessary to avoid a lack of suffering, in terms of a consensus on animal welfare assessment, there has been little attempt to define the levels of freedom that are desirable or the adverse consequences of not providing such freedoms. These different frameworks or concepts on animal welfare lead to the use of different methodologies to assess an animal's

welfare and therefore it is useful to briefly consider these concepts and their rationale, before considering these methodologies.

Three conceptual frameworks dominate interpretation of animal welfare. These are

- Biological functioning
- Mental functioning (or affective state) and
- Naturalness

3.1. Biological Functioning Concept

The biological functioning concept, equates poor welfare to difficult or inadequate adaptation (Broom, 1986; Hemsworth and Coleman, 2011). Broom (1986) defines the welfare of an animal as "its state as regards its attempts to cope with its environment". The "state as regards attempts to cope" refers to both (1) how much has to be done in order to cope with the environment and includes biological responses such as the functioning of body repair systems, immunological defences, physiological stress responses and a variety of behavioural responses and (2) the extent to which these coping attempts are succeeding. These behavioural and physiological responses include abnormal behaviours, such as stereotypies and redirected behaviours, and the stress response, respectively, while the success of the coping attempts are measured in terms of lack of biological costs, such as adverse effects on the animal's ability to grow, reproduce and remain healthy and injury-free (i.e., fitness effects). Thus the intensity of challenges from the animal's environment , including social and climatic stressors and environmental complexity, and challenges from disease including infections, traumatic injury, and poor nutrition, will be reflected in the magnitude of the biological responses utilised by the animal in its attempts to cope. As Broom (1986) recognises, there are two general types of indicators of poor welfare, one demonstrating that an individual has failed to cope with an environment, the other indicating the effort involved as the individual attempts to cope.

More detailed and recent accounts of the rationale of this concept are provided by Mellor et al. (2009) and Hemsworth and Coleman (2011).

This definition of Broom's (1986) is not dissimilar from the one recently endorsed by the 172 member countries of the OIE (2008): "Animal welfare means how an animal is coping with the conditions in which it lives. An animal is in a good state of welfare if (as indicated by scientific evidence) it is healthy, comfortable, well nourished, safe, able to express innate behaviour, and if it is not suffering from unpleasant states such as pain, fear, and distress. Good animal welfare requires disease prevention and veterinary treatment, appropriate shelter, management, nutrition, humane handling and humane slaughter/killing. Animal welfare refers to the state of the animal; the treatment that an animal receives is covered by other terms such as animal care, animal husbandry, and humane treatment."

Some have narrowly interpreted this biological functioning concept of animal welfare as one that equates an animal's welfare to attempts to cope with the environment, whether successful or not (Korte et al., 2007), however, others such as Moberg (2000) and Barnett (2003) emphasise that it is the consequences of these coping attempts that determine an animal's welfare rather than the responses per se. A key precept in this concept is that animals use a range of behavioural and physiological responses to assist them in coping with environmental conditions, and while biological regulation in response to environmental change is constantly occurring, adaptation is not always possible. When homeostasis fails, there is damage, disease or even death (Broom, 1986; Moberg,

2000; Barnett, 2003). Therefore, difficult or inadequate adaptation generates animal welfare problems.

Others have criticised this concept of animal welfare on the basis that it does not adequately include emotions or feelings. However, this would only be valid if emotions are independent of other biological processes but this is unlikely since the mental state of an animal is an integral component of its biological state (Dantzer and Mormede, 1983). Emotional responses are produced in the limbic system, which projects to several parts of the brain, including those involved in the initiation and maintenance of the stress response, thus explaining why emotional insults activate a stress response (Kaltas and Chrousos, 2007). Emotions are part of the body's regulatory system and together with a range of learning processes function to assist animals in avoiding potentially harmful situations or recognizing potentially beneficial situations (Cabanac, 1979).

In conclusion, how well an animal is coping with the challenges it faces will be reflected in the normality of its biological functioning and fitness, and severe risks to welfare will be associated with the most extreme coping attempts. Difficult or inadequate adaptation will affect the fitness of the animal through a range of long-lasting behavioural and neuroendocrine responses and thus the rationale underpinning this animal welfare concept of biological fitness is that difficult or inadequate adaptation generates welfare problems for animals. These behavioural and physiological responses include abnormal behaviours, such as stereotypies and redirected behaviours, and the stress responses including those involving both the sympathetic-adrenal-medullary and the hypothalamic-pituitary-adrenal axes, respectively, while the biological cost includes adverse effects on the animal's ability to grow, reproduce and remain healthy and injury-free.

A focus on biological functions during welfare assessment emphasizes

- Behaviour
- Homeostasis in the animal's physiology
- Health and disease
- Genetics, and
- The concept of interrelated costs and benefits between biological functions that influence adaptation, fitness and failure of the animal to cope with stressors that leads to pathology

3.2. Affective State Concept

The affective state or feelings-based concept, defines animal welfare in terms of emotions and emphasizes reductions in negative emotions, such as pain and fear and frustration, and increases in positive emotions such as comfort and pleasure (Duncan and Fraser, 1997). It should be recognized that there are numerous definitions of emotions in the literature often representing several disciplines. Denton et al. (2009) view primordial emotions as the subjective element of the instinctive behavioural patterns. These primordial emotions include thirst, hunger for air, hunger for food, pain, hunger for specific minerals, sexual arousal and orgasm, sensations accompanying impediment of visceral function (e.g., for micturition or defecation), desire for sleep after severe deprivation, and avoidance of change of body core temperatures, etc. (Denton, 2006). Denton et al. (2009) contrast the primordial emotions with another class of emotions which are most often fired by the distance receptors (exteroceptors)—the eyes, ears and the nose. These distance receptor evoked emotions, like rage, fear, hate, envy, happiness, playfulness, affection, anxiety, depression and disgust, are those to which the term emotion is most commonly applied.

Duncan (2004; 2005) has argued that animal welfare ultimately concerns animal feelings or emotions as follows. All living organisms have certain needs that have to be satisfied for the organism to survive, grow and reproduce and if these needs are not met, the organism will show symptoms of atrophy, ill-health and stress and may even die. Higher organisms (vertebrates and higher invertebrates) have evolved 'feelings' or subjective affective states that provide more flexible means for motivating behaviour to meet these needs. Thus the central argument is that although natural selection has shaped animals to maximize their reproductive success, this is achieved by proximate mechanisms involving affective states (pain, fear, separation distress, etc.) which motivate behaviours that can ultimately enhance fitness (Fraser, 2003).

Animal emotions have in the past been considered inaccessible to scientific investigation because they have been described as human subjective experiences or even as illusory concepts outside the realm of scientific inquiry (Panksepp, 1998). The difficulties in studying emotions as though they were objective states of bodily arousal are well recognized in the literature (Cacioppo et al., 1993). While each emotion may reflect a different pattern of arousal, the visceral response to many emotions is reasonably uniform in animals. Most animals react physiologically, at least in the short term, in essentially the same way whether the arousal is sexual, fear provoking or if there is the anticipation of play or food. It is obviously a major challenge to study and understand emotions in animals, although there have been some promising recent developments in the comparative study of emotions that show that there are many homologous neural systems involved in similar emotional functions in both humans and other mammals, and perhaps other vertebrates (LeDoux, 1996; Panksepp, 1998, 2005).

It is widely accepted in animal welfare science that good welfare is not simply the absence of negative experiences, but rather also requires the presence of positive experiences such as pleasure (Boissy et al., 2007; Mellor et al., 2009). While methods to assess pain and suffering have been developed, there is still no agreement on how to assess positive experiences (Boissy et al., 2007). However, preference tests have been used to identify resources and behaviour that might be important to hens (Cooper and Albentosa, 2003). Furthermore, measuring preferences of animals, using preference tests, aversion learning and behavioural demand testing (Dawkins 1980; Matthews and Ladewig 1994; Kirkdon and Pajor, 2006) has been used by scientists to assess animal welfare predominantly on the basis that these preferences are influenced by the animal's emotions.

A focus on how the animal feels (mental functioning/affective state) during welfare assessment recognizes that animals feel emotions; that emotional states are important to an animal's well being and that environmental, social and infectious conditions can induce negative emotional states that compromise welfare.

3.3. Natural or Normal Behaviour Concept

The third main concept of animal welfare, which is not often well-enunciated, promotes the principle that animals should be allowed to express their normal behaviour. For some this also implies that animals should be raised in 'natural' environments and allowed to behave in 'natural' ways.

The term abnormal behaviour in domestic animals invariably raises questions about what is normal (Mills, 2010), particularly when most behavioural differences between wild and domestic animals appear to be quantitative rather than qualitative in character, and best explained in differences in response thresholds (Price 2003). When considered as an aspect of the behaviour of an animal, abnormal behaviour is frequently defined as behaviour that is either atypical for the species, outside the normal behavioural pattern that has evolved in the natural habitats of the species or outside the range usually observed in the species in non-captive situations (Keeling and Jensen, 2005). In the

early literature, the view that animals should perform their full 'repertoire' of behaviour was very common, however there is broad agreement within science that it is often difficult to attribute actual suffering when the expression of certain behaviours is prevented or is absent when it would be expected to be present (Dawkins, 2003). Furthermore, as Fraser (2003a) notes, "Few scientists today would support the simple view that animal welfare depends on the animal carrying out all its natural behaviour in a natural environment because natural environments contain many hardships (harsh weather, predators), and natural behaviour includes many means of dealing with hardship (shivering, fleeing)."

The difficulty of deciding what constitutes the natural environment for domestic animals is brought in to focus when reviewing the history of the domestic hen as described by Appleby *et al.* (1992). The progenitor of the domestic fowl was the Red Jungle Fowl (*Gallus gallus*). It is a tropical species confined to forested areas and to thick vegetation. There are now two modern hybrids, the egg laying bird that reaches point of lay at 16-18 weeks of age at a body weight around 1.8-2.0 kg and that lays close to an egg a day, and the meat bird which reaches slaughter weight of about 2.0 kg as quickly as 5 weeks of age. What is the 'natural environment' of a young bird selected for meat production or an adult hen selected for egg laying, both of which are the same species, and following about 8000 years of selection for fighting capabilities and a hundred years of intense selection for production attributes? Is an outdoor area with relatively little structural diversity, except perhaps for some grass, a natural environment for a tropical species?

Thus the concept of 'natural' would need to be more specific before it could give guidance in assessing animal welfare, since generalizations may lead us astray and achieve the opposite of what is intended. Similarly, the 'natural behaviours' that are desirable or undesirable in terms of animal welfare require definition together with the rationale for their inclusion or exclusion. More recently the emphasis has been on behavioural indicators of poor coping such as fearfulness, aggression and stereotypies (EFSA, 2005), responses that are also utilized in the biological functioning-based concept of animal welfare.

Related to this notion of the importance of displaying normal behaviour is that of 'behavioural (or ethological) need'. The term 'behavioural need' appears to have been introduced into the scientific literature without any scientific evidence (Duncan, 1998). Dawkins (1990) and Fraser and Duncan (1998) suggested that the term 'behavioural need' refers to situations that elicit intense negative emotions and likely evolved for those behaviours in which an immediate action is necessary to cope with a threat to survival (e.g., escape from a predator) or reproductive fitness (e.g., nesting). In contrast other types of behaviour that can be performed when the opportunity arises (e.g., play, grooming) are more likely to be associated with positive emotional states. Duncan (1998) defined "behavioural needs" as behaviour patterns that are very strongly motivated, and, if they are not allowed expression, the animal's welfare may be jeopardized. However, any argument for impaired welfare due to restriction of these behaviours would be strengthened by supporting physiological measurement of frustration (Cooper and Albentosa, 2003), or evidence of decreased health or increased physiological stress (Duncan (2005).

A focus on natural behaviours in welfare assessment has been more strongly driven by philosophical precepts than by biological evidence. Defining natural behaviour and understanding the impact of the inability to perform them remains a major unanswered question in the assessment of animal welfare from the perspectives of biological function and affective states.

3.4. Scientific Uncertainty

These different concepts or views on animal welfare can lead scientists to use different criteria or methodology in assessing an animal's welfare. For short term animal welfare issues involving acute stress, such as painful husbandry procedures, there is considerable agreement on the need to assess animal welfare from a perspective of biological functioning (Mellor et al., 2000). However, for longer term issues, disagreement over these welfare concepts, especially when consequent interpretations conflict, often lead to debates concerning animal welfare and the varying interpretations (Fraser, 2003a,b).

This so-called 'scientific uncertainty' does not necessarily diminish the robustness of the research utilising methodologies or measurements arising from these views or concepts, but it does raise the question of the relatedness of these concepts (Barnett and Hemsworth, 2009). In other words, are situations in which an animal has to resort to the extreme coping attempts (i.e., challenges that may overwhelm an animal's capacity to adapt) associated with, or do they lead to, negative affective states and vice versa? In a similar context, is an inability to perform normal or 'natural' behaviours associated with extreme coping attempts and/or negative affective states? Therefore, if these concepts are related, are the resultant methodologies measuring the same adverse physiological and mental state(s) in the animal? Indeed many authors have raised the commonalities in these concepts (e.g., Fraser, 2003[b], 2008).

As suggested by Barnett and Hemsworth (2009), this conceptual convergence suggests a way forward in developing a broader consensus on the study of animal welfare by reducing both conceptual differences and consequently methodological differences in animal welfare science. The validity of the welfare criteria can be tested in several ways: first, with the finding that there are correlations between independent measures of different concepts of animal welfare; and second, with the finding that an intuitively aversive condition reduces animal welfare on the basis of the measures of different concepts of animal welfare. Therefore, research examining the validity of these concepts—and, in turn, methodologies—is necessary to understand the relationships between the concepts and indeed minimize the conceptual and methodological differences as discussed here. The development of a broader scientific consensus on welfare measures arising from this research should lead to the development of credible measures that can be incorporated into welfare assessment and screening tools in the field. It should be noted that there is indeed some evidence to support this conceptual convergence (Nicol et al., 2009; Stevens et al., 2009; Arnold and Matthews, 2010; Matthews and Bryant, 2011).

In the meantime, until science can broadly agree on the best methodology or methodologies to evaluate animal welfare, these approaches should guide current welfare research methodology. Using several of these approaches where the opportunity arises should also be utilised: for example as Widowski and Hemsworth (2008) recommend that, while studies of motivation can provide compelling evidence that the performance of some behaviour (or preference) may be important to the animal, additional evidence, particularly on occurrence of abnormal behaviour, stress physiology and health, are necessary to provide a more comprehensive assessment of the impact on animal welfare. Furthermore, the basis of the methodology used by scientists to assess animal welfare should routinely be provided so that individuals using science in their decision-making appreciate both the rationale for the methodology and its limitations (Fraser, 2003b; Sandoe et al., 2004).

There is evidence of a convergence of the three conceptual frameworks for assessment of animal welfare. Nonetheless, intrinsic scientific uncertainty remains in these approaches. This uncertainty does not diminish the value of scientific approaches to the assessment of animal welfare.

3.5. Interplay between Welfare and Ethics

We can see from the above discussion that animal welfare is at least, in part, a conceptual construct developed by humans to protect animals within our care and responsibility. No doubt it has also been developed to protect some of the moral sensibilities of humans. Importantly, animal welfare does not stand independent of ethics. The boundaries to what is considered to lie within the domain of animal welfare are influenced by religious and traditional understandings and change with emerging philosophical and biological knowledge. Interpretations of what within these boundaries constitutes good welfare come under similar influences.

Whist there are some (Broom, 1996) who suggest the welfare assessment and ethics are juxtaposed, others (Sandoe et al., 2003) would argue they are inextricably linked and this should be made more transparent by the scientific community. According to Sandoe (2011), the application of ethics occurs at two levels. Firstly, it occurs during the derivation or establishment of the scientific conclusions. For example, in the investigation of a specific welfare issue (e.g. stocking density, transport duration), the outcomes are rarely black and white. Quite often this occurs because there aren't clearly defined thresholds indicative of acceptable and unacceptable welfare in the measured responses. Therefore, when drawing a conclusion about a minimum stocking density for example, it is ultimately a subjective decision based on the consideration of the facts and an ethical context. The second level occurs during the evaluation and processing of scientific evidence. Any judgement about whether to accept or not accept scientific evidence pertaining to the welfare status of an animal or a production system or practice will depend on the individual's ethical perspectives and values. What may be deemed reasonable by scientists may not be to others simply because of their different ethical points of view.

Ethicists recognise many concepts that can be used to help judge what is right and wrong conduct. These concepts such as duty, rights, utilitarianism, and consequentialism provide additional influences on the interpretation of welfare states. Most participants, including scientists, engaged in assessment and interpretation of welfare are not schooled in ethics and do not knowingly bring formal ethical principals into their consideration of welfare (Mather, 2011; Veissier et al., 2011). Nonetheless, the nascent or unformulated ethical viewpoints of those participating in discussions on welfare can engage values that lead to divergent and strongly contested conclusions. No simple means to reconcile these divergent viewpoints is apparent, although procedures such as the Ethical Matrix and the ethical assessment process of Campbell are being developed for application in animal welfare (Croney and Anthony, 2010). A more detailed discussion of ethics is beyond the scope of this review.

Ethical values influence the choice of measures used during the assessment of animal welfare and the interpretation of data. Progression from collection of data to its interpretation and subsequent deliberation of ethical questions is usually a multistep, iterative process.

4. Welfare Assessment Measures Used in Science

In science, a broad range of measures have been applied to assess the welfare of experimental animals. The actual choice of measures or methodologies will vary depending on the experimental design, type of experimental challenge/treatment and livestock species. The measurement categories, and examples of specific measures within each, are shown in Table 1.

In addition to the application of existing welfare measures in livestock production research, considerable research effort has also been expended developing novel measures and approaches to

assessing changes in biological function and affective state. The latter has received significant attention over the last 20-30 years and this is discussed in more detail below.

A broad range of measures have been applied to assess the welfare of animals. Choice of measures or methodologies will be influenced by the type of experimental challenge/treatment and livestock species. Measurements can be broadly categorised as those relating to the animal's behaviour, productivity, health, physiology, affective state, environment/resources and genetics/genotype.

Table 1: A selection of commonly used measures from which welfare status of animals is inferred. See notes for explanation of terms. The quantification of the value and interpretation of these measures shown in the table is quite subjective and is intended to provide a guide to the challenges faced with using the measures

Category	Examples	Extent to which measure integrates impact of multiple stressors over an extended period	Validity as a welfare measure: strength of association with welfare issues and sensitivity to non welfare effects	Technical robustness of measure – repeatability across operators	Is the measure an indicator of good welfare, poor welfare or both?		Practical application				Limitations	Knowledge gaps
					Good welfare	Poor welfare	On-farm	QA audit	Vet-clinic	R&D		
Explanatory notes	**1**	**2**	**3**	**4**	**5**		**6**					
Behaviour												
Ethograms	Behavioural schedules, "natural behaviours"	+	+	++	+	++	?			✓	Absence of expression doesn't mean welfare is compromised	Impacts of lack of opportunity to express behaviours
Normal/abnormal	Pain related Injurious behavior, thermoregulation	+ +++ ++	+++ +++	++ +++		+++ +++	✓ ✓	✓ ✓	✓ ✓	✓ ✓	Can vary with type of pain. Species specific	Comparisons across types of pain within species and between species
Social behaviours	Agonistic, affiliative, voluntary flocking or isolation	++	+++	+++	+++	+++	✓		✓	✓	Species specific, influenced by age structure, gender	Impacts of lack of opportunity to express behaviours
Human-animal relationship?	Fear of humans, habituation to management, temperament traits	++	+++	++	+	+++			✓	✓	Negative experience prior to purchase not necessarily remediable,	Non-genetic methods for modifying temperament

											genetic influences	
Productivity												
Growth/Composition	ADG, growth targets, body condition score, feed conversion efficiency	+++	++	+++		+	✓	✓	✓	✓	Genetic variation (within species) influences target values multiple influences	
Reproduction	Fertility, fecundity, age at puberty, Days open, return to service, parity number success failure	+++	++	++	+	++	✓	✓	✓	✓	Genetic variation (within species) influences target values More difficult to measure in extensive systems, multiple influences	
Product yield	Milk Wool Egg Meat	+ +++ + +++	++ + + +	+++		++	✓	✓	✓	✓	Genetic effects, species specific multiple influences	
Product quality	Milk Wool strength Egg Meat	+ + + +	+ +++ + ++	+++		++	✓	✓	✓	✓	Genetic effects, species specific multiple influences	
Culling variables	Percentage, age, reason for culling	+++	++	+++	+	++	✓	✓	✓		Genetics influences, enterprise objectives multiple influences	
Health												
Infectious disease status	Mastitis, footrot, bovine respiratory	+	+++	+++	++	++	✓	✓	✓	✓		

	disease complex, parasites											
Metabolic disease status	Ketosis, acidosis	+	+++	+++		+	✓		✓	✓		
Developmental disease status	Joint and long bone deformities	+++	+++	+++	++	++	✓	✓	✓	✓		
Physical state of animal	Clinical signs including: coat (pelage), demeanor, eyes, reflexes food intake, faeces, urine	+	++	+++	++	+++	✓	✓	✓	✓		
Physical injury	Lameness, cannibalism	+++	+++	+++	+	+	✓	✓	✓	✓		
Mortality		+++	+++	+++	+	+	✓	✓	✓	✓		
Physiology												
Sympatho/ adrenomedullary system	Catecholamines, Heart function variables, vagal tone	+	+	++	+	+			✓	✓	Most samples need lab analysis	
Neuroendocrine	CRH, ACTH, Cortisteroids, prolactin, oxytocin, vasopressin, etc	+	+	++		+				✓		
Immune function	Hematology, antigen responsiveness, acute phase proteins	+	+	++		+			✓	✓		
Metabolic status	Ketosis, acidosis, mineral (micronutrient) status	+	++	+++		+			✓	✓		
Neurotransmitters	Dopamine, glutamic acid, endorphins, neuropeptides	+	+	++	+	+				✓		
Affective state												
Behavioural demand		+	++	+		+				✓		Importance of opportunity to experience negative affective states
Cognitive bias	Judgement bias, Attention bias	+	++ or +?	+	+?	+				✓		
Preferences		+	++	++		+				✓		
Qualitative behavioural assessment		+	++	+++		+++				✓		
Environment/resource												
Social	Enforced isolation, group size and structure	++	+++	+		+	✓	✓		✓		

Climatic	Temp, humidity, atmospheric pressure	+	+	+++		+	✓		✓	✓		
Physical	Housing, sun shade, wind, bedding, room to move, cleanliness predation	+++	+++	+++		+++	✓	✓	✓	✓		
Nutrition	Quality, abundance,	+	+	++		+	✓		✓	✓		
Stockmanship	Attitude to animals, use of goads, husbandry skills, records and review processes, participation in QA systems	+++	+++	++	+++	+++	✓	✓	✓	✓		
Genetic variables												
Breeding values	Birth weight Disease resistance Flight time	na	++	++			✓			✓		Genetic and phenotypic costs and benefits with other welfare and production traits Multiple influences?
Molecular markers	Polledness Disease resistance Fear?	na	++	++			✓			✓		

Notes

1. A non-exhaustive list of examples of each category of measures. Individual measurements can inform welfare interpretations in more than one of the 3 major domains: biological functioning, mental states, and naturalness discussed above.
2. Some measures are highly labile and change quickly with the circumstance of the animal, whereas others represent the cumulative effect of extended exposure to stressors. Measures which integrate extended exposure to stressors have more + signs
3. Validity as a welfare measures scores the strength of association of the measure with compromised welfare (+++). Measures with a low score are more susceptible to perturbation by events that don't necessarily compromise welfare can perturb.
4. Technical robustness scores the extend of standardization of the measure and ease of repeatability across operators
5. Some measures can indicate that welfare of the animal is compromised but absence of the measure does not necessarily indicate that the animal is in a good welfare state. An example is lameness. These types of measures are scored under the column as indicators of poor welfare but are not scored as indicators of good welfare. Scores tend to be low in the good welfare column because few measures have amplitude in the direction of good welfare to indicate degrees of very good, or extremely good welfare.
6. Practicality for use by farmers (on-farm) in QA audit systems, in veterinary investigations of animal health, and in research settings.

4. 1. Changes in Biological Functioning

As discussed in Section 3.1, this approach to assessing the welfare of an animal focuses on the normality of its biological functioning and consequent fitness. Difficult or inadequate adaptation will affect the fitness of the animal through a range of long-lasting behavioural and neuroendocrine responses. This approach has been used by scientists to assess the effects of housing, husbandry and handling on animal welfare. For example, a broad examination of the behavioural, physiological, health and fitness responses in handling studies, particularly in pigs and poultry, have generally shown that negative or aversive handling, imposed briefly but regularly, will increase fear of humans and reduce growth, feed conversion efficiency, reproduction and health of these animals (see Waiblinger et al., 2006; Hemsworth and Coleman, 2011). A chronic stress response has been implicated in these effects on productivity since in many of the pig handling studies (see Hemsworth and Coleman, 2011), handling treatments which resulted in high fear levels also produced either a sustained elevation in the basal free cortisol concentrations or an enlargement of the adrenal glands. Studies examining surgical husbandry procedures have also used a broad examination of the behavioural, physiological, health and fitness responses to study animal welfare (Mellor et al., 2000; Hemsworth et al., 2009; Colditz et al., 2010).

Studying biological function thus provides the opportunity to identify disturbances arising both internally (deviations from internal functional states, e.g., deviation from homeostasis) and externally deviations from expectations, e.g., emotions) but appear to provide little opportunity to identify positive experiences.

There are a large number of parameters that can be measured that reflect changes in biological functioning in response to challenges that can arise during animal production processes. These have been reviewed elsewhere (e.g. Moberg and Mench 2000) and continue to evolve as new understandings of physiological and behavioural responses develops. For example, Wang *et al*. (2004) discovered a new nociceptive signalling pathway, and suggest that the chemical mediator superoxide could be used as a novel indicator of pain. Nonetheless, much remains to be discovered by research on molecular markers of stress (Gornati *et al.* 2005). In addition, considerable efforts are under way investigating the potential for measures of biological functioning to reflect mental health in animals (e.g. Yeates and Main 2007); and new technical developments such as infra-red thermography will enable increasingly sophisticated measurement on animals without the confounding effects of stress arising from the measurement process itself (Stokes et al., 2012). The main methodological issues with the measurement of biological functioning are:

- The lack of an agreed procedure for amalgamating different measures into a global welfare index, and
- Determining the levels of biological functioning that match with different levels of welfare.

Notwithstanding the different ethical viewpoints amongst the various stakeholders on the acceptability of specific husbandry practices, the lack of clear scientific guidelines for integrating measures and rating practices (Fraser 1995) has lead to divergent recommendations from within the scientific community on the acceptability of various procedures. For example, two different reviews of the available scientific evidence on the housing of sows came to different conclusions. Barnett *et al.* (2001) emphasised biological functioning and corresponding decreases in fitness in assessing animal welfare (e.g., criteria such as behaviour (aggression), stress (cortisol), health, immunology, reproduction, injuries, growth rate and nitrogen balance) and concluded that "On balance, it would appear that both individual and group housing can meet the welfare requirements of pigs.". The second review, by von Borell *et al.* (1997), emphasised the importance of both affective states and the opportunity to carry out natural behaviour in assessing animal welfare (e.g., high levels of abnormal behaviour and inability to perform some natural

behaviours) and concluded that "Since overall welfare appears to be better when sows are not confined throughout gestation, sows should preferably be kept in groups." Underlying differences in ethical values brought to bear by the authors during interpretation of the results appear to have led, at least in part, to these divergent conclusions.

Extending Broom's definition of animal welfare (Broom, 1986), Moberg (2000) has proposed, in our view correctly, that the key to the development of an index of welfare based on biological functioning is the measurement of the biological cost of challenges (at least for harmful challenges). Such an index will be underpinned by measures of endocrine, or other physiological and behavioural responses to challenges, but as none of these provide definitive endpoints that reflect the biological cost to the animal, nor are easily integrated (Fraser 1995), the index is unlikely to be based on them alone. Further, different constellations of responses are seen depending on the nature of the challenge and other genetic and life history experiences (Jarvis *et al.* 2006). Thus, Moberg (2000) has proposed a model based on pre-pathological conditions as quantifiable measures of biological cost. Such conditions include disruption of ovulation, abnormal growth and abnormal behaviours as well as sub-clinical or clinical disease. While measures are more quantifiable and more integrated than individual responses, such as changes in adrenal output of endocrine hormones, this proposition has the disadvantages that:

- there is no common index to rate different practices but this challenge is common to many other approaches ;
- it restricts concern about welfare issues to insults; and,
- there is little possibility for detecting positive experiences.

There are a range of alternative possibilities for developing an overall index for welfare assessment and these are presented below.

Numerous measures have been successfully applied to assess changes in biological function. Key limitations include:

- The lack of an agreed methodology for amalgamating different measures into a global welf index, and
- Determining the levels of biological functioning that match with different levels of welfare.

4.2. Assessment of Negative and Positive Mental States in Livestock

While the subjective experiences of animals cannot be measured directly, there is a consensus, as embodied in the Five Freedoms, that animals have the capacity to suffer and experience negative affective states. Further, there is a common belief amongst many citizens that animals have the ability to experience positive affective states (Kjaernes *et al*. 2007). Animals respond peripherally to challenges with a wide variety of responses that result in activation of some or all of the following biological process: immune system, hypothalamic-pituitary-adrenal axis, motor and other behavioural activities (e.g. Appleby 2011). Interacting with these peripheral responses is a range of neurophysiological changes, which influence both peripheral physiological responses and mental/cognitive processes. Cognitive elements include the animal's perception of its own state (e.g. the aversiveness of the events/feelings). Most of the scientific research on animal welfare has aimed at assessing the effects of putatively stressful or harmful events on animals' wellbeing, although there is increasing interest in understanding if other aspects of the production environment can contribute to positive welfare (Boissy *et al.* 2007; Yeates and Main 2007). Measurement of the harmful effects has focussed on understanding the physiological (peripheral, and to a lesser degree, central) and behavioural responses to stressors.

There is growing recognition by the scientific community of the relevance of an animal's subjective experiences to key stakeholders and there has been steady progress in the development of techniques

to more directly access the perceptions and feelings of animals (reviewed by Dawkins, 2006). These will be discussed in detail as they present the major opportunity to align scientific methodologies with community views (Hogan and Fisher).

Preference Testing

One of the initial methodologies to be applied in this context was based on preference testing. The best example is the use of a Y-maze that allows a choice between access to two different resources. This approach has been used extensively to provide information about specific features in the animal environment such as flooring (Hughes and Black, 1973; Hutson, 1981), restraint methods (e.g. Pollard et al., 1994), handling treatments (Rushen, 1986) and ramp design (Phillips et al., 1988), with the overriding objective of optimising the environment for animals.

While the consistent choice or preference of one resource over another or others indicates the animal's relative preference, some have argued that in addition to establishing what an animal prefers, it is important to understand the strength of the preference (Dawkins, 1983; Matthews and Ladewig, 1994). To address the question of the strength of an animal's preference, experiments have incorporated varying levels of cost (e.g., work effort, time and relinquishing a desirable resource) associated with gaining access to a resource or avoiding aversive stimulation (refer to the subsequent section on Behavioural demand). For example, Dawkins (1983) varied the price paid for access to litter by increasing the duration of feed withdrawal before the test. She found that although hens preferred litter to wire floors, their preference was not strong enough to outweigh the attraction of food and concluded that in both experiments there was no evidence that hens regarded litter as a necessity.

These initial preference studies stimulated considerable debate on conceptual and methodological difficulties (Dawkins, 1977; Duncan, 1978) and Fraser and Matthews (1997) concluded that the usefulness of preference tests to answer questions about animal welfare is limited by three main issues. First, these tests should adequately reflect the animal's preference, second, they need to establish how strongly an animal prefers a chosen option, avoids a non preferred one or is motivated to perform a behaviour, and third, preferences may not correspond to welfare if the choices fall outside the animal's sensory, cognitive and affective capacity or if the animal is required to chose between short- and long-term benefits.

Expanding on these limitations, firstly, preference tests measure an animal's choice behaviour at a point in time and such measurements run the risk of failing to account for interactions of different motivational states which may influence the behaviour of the animal over time (Hutson, 1984). Furthermore, this short term choice may reflect the animal's proximate (immediate or present) requirements, rather than the animal's ultimate requirements or those necessary for survival, growth and reproduction (Lawrence and Illius, 1997). Clearly preference tests therefore need to be comprehensive enough to identify these sources of variation (Fraser and Matthews, 1997).

Secondly, preferences may vary with familiarity (Phillips et al., 1991, 1996) and thus prior experience can be controlled in preference tests by using naïve animals, familiarizing the animals with the resources prior to testing or, as Fraser and Matthews (1997) suggest, prolonged testing. Furthermore, since preferences for specific resources may also be affected by the context in which the animals are tested (e.g., the social environment at the time of testing), the context in which the animals are studied therefore should relate to the commercial conditions in which the experimental question is directed (Dawkins, 2003).

Thirdly, Fraser and Matthews (1997) recognize that limitations in using preference tests arise when animals are exposed to potential dangers or benefits that are beyond their sensory or affective capacity or if the choice requires a level or type of cognitive ability that the animal does not possess. They suggest that the best safeguard is to base preference tests on the types of choices that the species arguably evolved the capacity to make and that the individual animals are accustomed to making.

In concluding on animal preferences, clarifying the conceptual link between animal preferences and animal welfare is an issue for some. The individual's concept of animal welfare clearly underscores the methodology used to judge or measure animal welfare. However, as commented by a number of authors (e.g. Fraser and Matthews, 1997), preference research should be integrated with other measures used in animal welfare research. Furthermore, Widowski and Hemsworth (2008) recommend that, while studies of motivation can provide compelling evidence that the performance of some behaviour (or preference) may be important to the animal, additional evidence, particularly on occurrence of abnormal behaviour, stress physiology and health, are necessary to provide a more comprehensive assessment of the impact of restriction on animal welfare.

Behavioural Demand

The importance of resources for animals can also be derived from measures of demand elasticity (Dawkins 1983). Consequently, 'behavioural demand' studies, using operant conditioning techniques in which the animal must learn to perform a response, such as pecking at a key or pushing through a weighted door, to gain access to a resource, have been used to study the animal's level of motivation to access or avoid the situation being tested. The strength of the motivation provides a quantitative measure of how much it matters to the animal. One methodology, derived from the theory of behavioural economics has proved helpful in identifying appropriate quantitative measures (Dawkins 1990; Lea 1978). Typically, in a behavioural economic framework animals are required to work for a resource, and the quantity of the resource obtained as the work requirement ("price') is increased is measured. The generic function (demand curve) describing the change in total quantity of the resource acquired as the price increases is positively decelerating (Hursh and Winger 1995) and takes the form:

$$\ln Q = \ln(L) + b[\ln(P)] - a(P)$$

Where,

Q is the measured consumption of the resource,

P is the price for a unit of the resource, and

L, b and a are parameters characterising the initial level of the curve at minimal price, the corresponding slope at minimal price, and the acceleration or increase in slope with increases in price, respectively.

ln natural logarithm

Elasticity (b – a(P)) is the point slope of this function and is a linear function of price. The price at maximal work (P_{max}) is calculated as (1+b)/a and occurs when elasticity takes the value -1. P_{max} can be conceptualised as the sensitivity of work output to environmental constraints and costs. The maximal work at P_{max} is O_{max}. O_{max} can be conceptualised as the level of resource seeking. The various curve parameters have been used in different ways to quantify animal perception, and there is some debate about which is the most useful (Kirkden et al., 2003; Kirkden and Pajor 2006*a*). Recent evidence presented by Verbeek et al (2012b) and Madden *et al.* (2007) suggest that O_{max} may be the best measure for quantifying the subjective experiences of animals. They reported that those resources (e.g. food or pharmacological agents) sustaining higher O_{max} values were the ones that more effectively met the animal's needs. O_{max} is equivalent to the area under the demand curve bounded by Pmax and the

corresponding level of consumption (Hursh and Winger 1995). In economic terms, O_{max} belongs to a class of measures known as consumer surplus, which Kirkden et al (2003) argue is the best measure of motivational strength or resource value. Resources with inelastic demand (elasticity values less than -1) have also been reported to reflect a strong need (Matthews and Ladewig 1994).

The validity of the behavioural economic approach is supported by other studies demonstrating that biological functioning is impaired in animals that are not able to access resources that are subjectively rated as very important. Mason et al. (2001) identified food and access to a water bath as needs, and that preventing access to one or other resulted in similar and elevated levels of stress hormone (cortisol) concentrations in the urine. Further, rest is rated as highly as food by dairy cattle (Munksgaard et al. 2005; Matthews et al., 2006), and reducing the amount of rest obtained each day by about 50% adversely effects the physical functioning of the animal in a variety of ways (e.g. altered hypothalamic-pituitary-adrenal axis regulation, Fisher et al., 2001; reduced growth, Fisher et al,. 2003).

Cognitive Bias Testing

Measures of cognitive bias are another way that has been proposed to assess affective state (both positive and negative) in animals (Harding et al., 2004). This methodology has been developed in human studies, where anxious or depressed individuals typically interpret ambiguous stimuli more pessimistically compared with non-depressed controls (Mathews et al., 1995). In the animal studies, the subjects are exposed to different treatments presumed to induce different affective states. By analogy with the human studies, animals are trained to respond to cues differentially associated with rewarded and unrewarded (or punishing) events. Typically, the rewarded and unrewarded cues are selected from a single sensory modality e.g. visual, auditory, tactile, and spatial. During testing for cognitive bias, stimuli intermediate to the training cues are presented and the animals' responses are measured. If the responses to the ambiguous stimuli are inhibited (e.g. slower) than the controls, then this is called a negative cognitive bias and is said to reflect a negative mood state induced by the treatment. Alternatively, if the animal's responses are, for example, quicker to the ambiguous stimuli, then this is called a positive cognitive bias and said to reflect a positive mood as a result of the treatment. There are growing number of studies, with several different species including rats, dogs, pigs, sheep and starlings, where a negative cognitive bias has been reported (see Mendl et al., 2009 for a review; Doyle et al., 2011a) and a number of studies beginning to report evidence of positive cognitive biases (e.g. sheep, Doyle et al., 2010; pigs, Douglas et al., 2012). However, there are a number of findings in cognitive bias studies which suggest that the interpretation of the data is not straightforward, at least in terms of ascribing states of positive and negative mood to the animals utilised. For example, treatments designed to induce different mood states sometimes produce no differences in cognitive bias (e.g. Burman et al., (2011) for supposedly positive states; Weichman et al. (2012) for supposedly negative states). Further, mismatches between the expected effect of treatment and the measures of cognitive bias have been reported. For example, positive bias has been reported following imposition of supposedly negative treatments like restraint and isolation in sheep (Doyle et al., 2010), and negative bias has been reported even though independent measures showed that there was no difference between treatments in the emotional state of the animals (Doyle et al., 2011b). As yet, the methodology is primarily a tool for understanding mental experiences of animals in experimental settings; if/when it becomes fully validated as a measure of affective state then practical measures for use in the field will be required.

Qualitative Behavioural Assessment (QBA)

Recent research has demonstrated the potential of the qualitative assessment of animal behavioural expressions as a valid scientific tool for the integration of different approaches to animal welfare (Rousing & Wemelsfelder, 2006; Defra, 2006; Stockman 2010; Rutherford 2012; Wickham 2012).

Qualitative Behavioural Assessment (QBA) is a whole animal approach that relies on the ability of human observers to integrate subtle information about animals' behaviour and body language and the animals' context to provide a valid measure of the animals' affective state. In other words, it describes not 'what' the animals do, but 'how' they do what they do.

These studies apply Free-Choice-Profiling methodology (FCP), and the associated statistical approach of Generalised Procrustes Analysis (GPA), to the qualitative assessment of animal behaviour. FCP elicits spontaneous descriptors of animal behaviour (e.g. "calm", "confident", "anxious"), and the level of consensus in those assessments among a group of observers can be calculated with GPA. Such descriptors have an expressive connotation that is relevant to how the animal perceives its immediate environment. This is preferred over other methods that focus on separate demarcated aspects of an animal's response such as when observers may be directed to make judgements based on pre-determined lists of indicators or body postures thought to reflect an animal's experience (e.g. flattened ears signal fear).

A number of QBA studies in pigs and other species have shown good internal validity (i.e. high levels of inter- and intra-observer reliability and repeatability) (Wemelsfelder 2001, 2009, Rousing and Wemelsfelder 2006, Walker 2010). Furthermore, external validity of QBA has been shown through correlations with quantitative behavioural measures (Napolitano et al., 2008; Minero et al., 2009 Rutherford 2012) and physiological indicators of stress in cattle (Stockman 2011) and sheep (Wickham 2012). Importantly, a recent paper further strengthens the biological validity of QBA as observer judgements were shown to be sensitive to the altered emotional state in pigs achieved through pharmacological intervention with an anti-anxiety drug (Rutherford 2012). Thus, this strongly supports the notion that QBA can be used as an outcome measure of emotionality (the affective state) in animals. A recent review of methodologies that might be used to assess positive welfare states in cattle concluded that QBA was 'the most promising assessment methodology' (Napolitano et al 2009). The UK Farm Animal Welfare Council report (FAWC 2009) has similarly indicated the important role that QBA could play in assessing the consideration of positive welfare states in animals.

A strong point of qualitative methodologies is that, given their integrative nature, they are sensitive to the context in which the observation is made. As many animal welfare studies seek to compare how animals cope in various contexts, there is a risk that contextual bias may occur if observers were to compare an environment they considered morally 'good' with one considered 'poor' (e.g. an enriched versus a barren environment). Wemelsfelder (2009) investigated how the perceived environmental background affects observers' assessments of pigs by comparing observers' judgements when the same subjects were viewed against an indoor and outdoor background by digital modification. High correlations were found between the pig scores viewed in both settings indicating that QBA is sensitive to context but this sensitivity does not weaken the reliability of such assessment.

Some concern for qualitative methodologies remains as to whether cultural backgrounds of observers and different levels of experience with animal behaviour observations may affect the reliability of observers. Napolitano et al (2012) supports previous QBA studies that have compared assessments from groups of observers with different nationalities and cultural and experiential backgrounds and shows reliable, high levels of inter-observer agreement between groups.

QBA represents a component of the multi-criteria Welfare Quality Project, a program designed to develop reliable on farm monitoring systems across the European Union. Further discussion of the Welfare Quality Project occurs in a subsequent section of this review. Hence, the studies described

above suggest that QBA has the potential to integrate other scientific measurements of animal welfare and to be used as a practical tool for on-farm welfare assessment and surveillance.

Assessment of Positive States

Consideration of positive welfare implies that good welfare is not just about the elimination of poor welfare but also includes aspects such as positive affective state (Yeates and Main 2007). While positive welfare has long been viewed by many citizens as an important aspect of good livestock husbandry, it has only recently become subject to critical scientific analysis. Methodologies to assess the capacity of livestock to experience positive mental states have been reviewed (e.g. Boissy et al. 2007; Yeates and Main, 2007; Mendl et al., 2010). While there are no unambiguous measures of positive states in livestock, there are a number of promising avenues under development. Fundamental neurobiological studies of feeding indicate that there are distinctive neural systems (and neurotransmitters) associated with the affective states (e.g. pleasure of having attained a reward, called "liking") and the motivation to obtain the reward ("wanting") (Berridge 1996; Berridge 2003). Wanting and liking are functionally and neurologically inter-related (Berridge and Robinson 2003). Thus, behavioural and physiological indicators of the states of liking and wanting would be useful candidates for measures of positive states. The demand function methodologies outlined earlier, and positive anticipatory behaviour (van der Harst et al 2003) are useful procedures for quantifying wanting. Methods for measuring "liking" include behaviours such as facial expressions and vocalisations, changes in cognitive functioning, immune and sympathetic nervous system parameters, and brain imaging techniques (for more detail see reviews by Boissy et al. (2007) and Yeates and Main (2007)). Interestingly, approach-avoidance behaviour (rather than the more usual measures of heart rate variability and high frequency vocalisations) has been identified as one of the best ways to assess the emotional valence of a situation in pigs (Imfeld-Mueller et al., 2011).

Indicators for positive welfare states that are suitable for on farm assessment have not been described; however appropriate environmental stimulation would favour good welfare and allow animals the opportunity to exhibit certain behaviours. Promising measures of positive indicators of cattle welfare include play behaviour which in calves is mainly expressed as locomotor (bucking and trotting) and social activities (rubbing and butting heads, play fighting), and social licking behaviour (Napolitano 2007[b]). A prominent aspect of the social behaviour of cattle is that these animals are gregarious and under semi natural or extensive conditions, their behaviour is highly synchronised. Thus a high degree of synchronisation of behaviours within the herd may indicate a positive welfare state (Metz 1983). As such behaviour may only occur at specific times of the day and between certain sub groups, making an instantaneous scan sampling technique is problematic (Napolitano 2007[b]). Hence the low feasibility (time consuming) and a lack of research into the reliability of play and social behaviours, means these types of measures, are not suitable for easy implementation into an assessment scheme.

Self Administration of Analgesics

A novel approach based on the measures of an animal's readiness to self-medicate analgesics, is particularly applicable to quantifying an animal's subjective experience of painful events. Danbury et al. (2000) trained broilers to discriminate between different sources of feed (with or without an analgesic). Lame birds selected significantly more analgesic feed than sound birds, and as the severity of the lameness increased, lame birds consumed a significantly higher proportion of the drugged feed. Thus, the severity of pain can be assessed from the measures of the amount of analgesic an animal will consume. This approach provides the most direct method for quantifying subjective experiences of pain and should be used much more widely in welfare assessment.

The combination of the behavioural economic and analgesic consumption methodologies, in which demand functions for access to analgesics by lame animals are determined, could provide additional

explanatory power regarding chronic pain. This "treatment demand" methodology would provide a more quantitative framework, and allow assessment of the animal's experiences without the potential confounding effects of large differences in drug/food consumption between treatments.

A variation on the treatment-demand methodology could be developed for quantifying other potentially unpleasant health conditions e.g. gastro-intestinal malaise, where animals could work for access to antacids or other treatments. With this range of methodologies it should be possible to extend the analysis of animals' experiences to a much wider range of potentially unpleasant challenges that has hitherto been possible or undertaken.

There has been steady progress in the development of experimental methods to assess affective states in animals. Preference testing and behavioural demand have been applied most frequently in this context. There are a number of promising behavioural and neurophysiological methodologies currently being evaluated. The ultimate utility of these new methods will be underpinned by their external validity (i.e. correlations with other independent measures of welfare state). QBA offers the most promise for the assessment of affective states on-farm.

4.3. Integration of Response Measures - Biological Function and Affective State

There are a number of different generic approaches that can be applied across species for integrating and linking the various societal/scientific aspects of the animal welfare debate, and which can be used to establish valid, practical indicators of welfare.

Using a cold challenge model with dairy cattle, Matthews and Bryant (2011) have demonstrated there is a good match between measures of affective state (as revealed by the animal making trade-offs between two highly valued but mutually-exclusive choices (shelter and rest)) and the level of biological functioning (as assessed by the energetic requirements to maintain thermal balance). Further, it was demonstrated that a practical measure, which reflects the degree of welfare challenge experienced on both the affective state and biological function dimensions, can be predicted from the degrees of cold below thermo-neutrality. A similar methodology has been used to assess the convergence between affective state and biological function in a heat stress model with dairy cattle (Arnold and Matthews, 2010). Similarly, Nicol et al. (2011) have shown that hens choose environments associated with lower stress (i.e. lower corticosterone levels and lower faecal water content). Preferred environments were also associated with behaviours such as less head shaking, self-scratching, standing alert and feeding, and more foraging, suggesting that responses such as these responses are indicators of good affective and physical states. With sheep and using a slightly different measure of affective state, a similar convergence between an animal's subjective experience of hunger and its requirements for energy has been demonstrated, both of which can be assessed practically with measures of body condition score (Verbeek et al., 2010, 2012[a], 2012[b]).) There is, thus, good evidence that this general approach can be used to establish and validate a range of quantitative, practical measures of welfare that reflect both the physical and emotional challenges faced by livestock.

Lateralisation

There is increasing evidence that degree of lateralisation (or handedness) can be used as a measure of both the affective and biological responses of animals to challenge (Matthews et al., 2012). Typically, animals show a preference to use one of a pair of bodily organs, or asymmetry in the use of medial organs, such as the tail (Quaranta et al., 2007). Strength of 'handedness' reflects (normal) right brain hemispheric specialisation for processing aversive emotional experiences (Kendrick, 2006) and a more responsive HPA axis (Westergaard et al, 2003). In sheep (Hernandez et al., 2009 [a]

[b], 2010) and other animals (Westergaard et al., 2003) stress inhibits handedness and emotionality. Measures of lateralisation have the advantage of perhaps being a relatively practical measure of welfare: the emotional and physiological responsiveness to an event can be assessed from such measures as the proportion of animals using their left versus right eye to evaluate the situation (Lippolis et al., 2005).

Chronobiological Measures

There is strong temporal (chronobiological) organisation of essential behavioural and physiological functions of animals and disruption to the patterning of activities can be used as a generic indicator of welfare status, for both health/disease diagnosis and non-health related conditions (Bergen, 2011). For diagnostic or welfare assessment purposes, the patterning of behaviour can be described quantitatively with a variety of mathematical functions (e.g. fractals, Rutherford et al., 2004) and compared with species-typical norms. While chronobiological measures require frequent monitoring of behaviour or physiology, recent developments in automated and remote monitoring technology, especially when combined with GIS (https://online.tugraz.at/tug_online/fdb_detail.ansicht?cvfanr=F28667&cvorgnr=37&sprache=2) offer unprecedented opportunities for practical animal welfare assessment.

Methodologies based on the integration between biological function and affective state measures have been successfully applied in livestock studies. They offer a more robust and compelling determination of welfare. Emerging methods such as the assessment of lateralisation offers promise in the context of unifying biological and affective responses in animals.

4.4. Can Welfare Measures Function as Diagnostic Tests?

Standardised measures or diagnostic tests are widely used for establishing the disease status of farm animals and humans. The tests are usually accredited for their diagnostic performance against standardised criteria including specificity, sensitivity, predictive value of a positive and predictive value of a negative as outlined in the figures below (Greiner and Gardner, 2000).

Diagnostic test	**Gold Standard**	
	Positive	**Negative**
Positive	**True Positive**	**False Positive**
Negative	**False negative**	**True negative**

Sensitivity and Specificity

- Sensitivity
  - the likelihood that a positive test result detects an animal that is positive for the gold standard (TP/(TP+FP)) × 100)
- Specificity
  - the likelihood that a negative test result detects an animal that is negative for the gold standard (TN/(TN + FP)) × 100)
- Predictive value of a positive
  - the likelihood that a positive test result comes from an animal that is positive for the gold standard (TP/(TP + FP)) × 100)
- Predictive value of a negative
  - the likelihood that a negative test result comes from an animal that is negative for the gold standard (TN/(TN + FN)) × 100)

For definitions see: Greiner and Gardner, Prev Vet Med 45:3

While diagnostic tests are typically used to identify infected versus not infected animals, some tests are also used to quantify the severity of an infection or a disease state (e.g. (Colditz and Le Jambre, 2008). In Australia, diagnostic tests for infectious disease are accredited through SCAHLS (http://www.scahls.org.au/) while best practice accreditation of laboratory standards for using diagnostic tests is provided though NATA (http://www.nata.asn.au/). In analogy with diagnostic tests of the severity of infection, measures of animal welfare can be considered to be tests aimed at detecting both the disease state of compromised welfare and the severity of that compromise. However, no individual measure described above (Table 1) approaches the level of sensitivity or specificity required for a reliable diagnostic test of compromised welfare. This shortcoming arises both from the diversity of physiological, behavioural and affective states that are considered either individually or in combination to constitute compromised welfare and the diversity of environmental, social, infectious and traumatic challenges that can perturb the individual measures without necessarily being considered to compromise welfare. Thus individual welfare measures fail in the standard performance criteria of sensitivity and specificity to diagnose compromised welfare. As a result, a suite of measures is routinely adopted for assessing welfare states. The aspiration for tests of animal welfare that achieve the sensitivity and specificity of diagnostic tests of infectious disease is unlikely to be met while ever the multiple dimensions in which welfare can be compromised remain clustered within the single common descriptor, animal welfare. A substantial challenge for measures of welfare is to define cut off points between positive and negative, good and bad, or acceptable and compromised welfare.

Two scientific approaches to define the boundaries between normal and abnormal are commonly employed. The first uses references values for the species, or for a subgroup of the species such as a breed when this differs from the species values (Lepherd et al., 2009). Reference values are established through measurement of a large sample of normal individuals from the population and

the distribution of normal values described by statistical parameters such as mean and 95% confidence intervals. Reference values permit interpretation of measures made on a single subject to determine whether the individual is clinically normal or abnormal. The second method uses statistical analysis of variables measured on the experimental or study groups under investigation. When animals within two or more groups are housed or managed in a way that lets contrasts be performed between the groups, statistical tests permit comparison to be made between the groups using the data from those groups in isolation of reference values for the species or breed. The important feature of the second approach is that statistical differences between values measured in the different groups or treatments can be detected that lie within the reference values for that variable within "normal" animals. Ascertaining the impact on welfare of treatments or management procedures that lie within the normal range remains a challenge for animal welfare science.

One approach to this challenge would be to establish a set of reference values for combinations of variables, taking into account interrelatedness of variables and the impact on the animal of having multiple variables simultaneously deviating from the mean.

Individual measures fail in the standard performance criteria of sensitivity and specificity to diagnose compromised welfare which is why a battery of measures is routinely employed. A key constraint is the lack of suitable reference values or critical thresholds for measures.

6. On-Farm Welfare Assessment Measures and Systems

There are a number of key drivers underpinning the need for welfare assessment systems on-farm. These include:

(i) Increasing societal concern about the treatment of animals and the need to further improve aspects of livestock production and to demonstrate these improvements to consumers.

(ii) Profitability in key markets will be influenced increasingly by the growing global trend of 'ethical consumerism' (Clarke *et al.* 2007) in which attributes such as animal welfare are becoming seen as components of food quality and hence influence consumer purchasing choices. Concomitantly, there will be an expansion in the demand for so-called "ethical foods" that satisfy the expectations of (relatively affluent) consumers seeking products that exceed regulated standards for animal welfare. This in-turn will require the establishment of a regulatory or accreditation framework that provides confidence to all stakeholders in the "truth in labeling" of the animal welfare and other ethical claims made for products.

(iii) The future growth in the human population will require an estimated 60% global increase in agricultural productivity by 2050 (OECD/FAO, 2012). To service this growth, there will need to be an intensification of animal production and farming, which is likely to occur at least in part through an expansion in the number and activity of large corporate and sovereign business entities. Furthermore, there will need to be further gains in the efficiencies of animal production. Collectively, these trends will continue to invoke societal and consumer concerns. Therefore, there will be a strong imperative to demonstrate that the welfare of livestock is optimised in future farming systems.

Despite the compelling need, no accepted comprehensive fully-validated system of welfare evaluation currently exists. Finding an acceptable welfare assessment system that is acceptable to all stakeholders is problematic. Various stakeholder groups tend to differ in the value frameworks under which they operate, although there are also differences within stakeholder groups depending on factors such as

culture and experience with farming practices (e.g. Evans and Miele 2007). Nevertheless, citizens/consumers are more likely to emphasise natural living, mental/emotional wellbeing and quality of life, while producers are more likely to align with the concept of physical fitness and biological functioning (e.g. Kjaernes et al., 2007; Matthews et al., 1994).

Further, it is obvious that different stakeholders will have different requirements for measurement methodologies and systems. Producers are primarily interested in indicators that will give them an early warning of impending conditions that will adversely affect the biological functioning and fitness of their livestock (Manning *et al*. 2007), regulatory agencies have a prime interest in compliance with minimum legal standards (related to the Five Freedoms in Australia), and marketing/retailers wish to see compliance with a comprehensive set of standards that reflect the views of their customers. Surveys undertaken in developed world markets indicate that consumers place high importance on emotional wellbeing, a state which is not currently readily assessable in animals (Kjaernes *et al.* 2007). Animal welfare advocacy groups/NGOs vary greatly in their needs for scientific assessment of welfare status. The moderate groups such as the Royal Society for the Protection of Animals (RSPCA) and Compassion in World Farming (CIWF) seek information to support their own welfare assurance schemes or campaigns along the lines of the Five Freedoms but with increasing emphasis on animal feelings. The more radical groups are heavily influenced by their beliefs and their actions may not take account of scientific information about welfare status. In practice, there is a mutual dependency between the key stakeholders (producers, marketers, consumers, advocates), so the information needs will need to reflect this.

Societal concerns, ethical consumerism and global food security are just some of the key drivers behind the need for farm animal welfare assessment systems. Despite a clear need, no one ideal system exists and this is highly unlikely given the diverse needs and interests of stakeholders.

5.1. An Ideal Welfare Assessment System

While the various stakeholders may, historically, have had different motives for developing and implementing welfare monitoring systems, the inter-dependencies between the different parts of the supply chain is leading to a convergence in the characteristics, requirements and application of welfare assessment. Simply stated, the assessment system must provide information that: enables producers to predict and maintain good standards of physical and mental wellbeing in their livestock; provides evidence that these standards have been achieved; and, demonstrates that the welfare outcomes are consistent with all three ethical frameworks.

Implicit in welfare assessment systems is the notion that it is possible to compare welfare standards between different production processes and systems, and rate each of them against some desired standard (e.g. Laywel, 2006). In other words, assessment procedures will require the application of methodologies to measure and rank the overall standards of welfare in different farming systems. It has been argued by some (e.g. Fraser 1995; Fraser et al 1997) that it is technically not possible to reduce welfare measurement to a single dimension and, therefore, it is not possible to compare overall standards of welfare. It is acknowledged that there are several major hurdles to be overcome in the development of overall welfare measurement systems e.g. ways to weight and integrate different welfare domains, but several research teams around the world are making progress in this area and alternative ways forward are detailed below.

Thus, an ideal assessment system would have the following features:

- The measures must be underpinned by scientific evidence demonstrating their validity (i.e. directly reflect the welfare states relevant to stakeholders). The measures will most likely be outcome-based, although some input measures (resources or management procedures) may also be relevant
- A framework for assessing the trade-offs or weightings between different welfare domains (e.g. nutrition, health)
- A methodology for integrating the weighted domains to develop an overall welfare index
- A method for identifying biologically relevant thresholds corresponding to different levels of welfare (e.g. minimum standard, gold standard)
- From a practical perspective, the measures must be technically feasible and implementable, reliable, reproducible, provide an early indication of impending welfare concerns so that livestock managers can take preventative action, and reflect the animal's welfare state over the assessment period.

For welfare monitoring on farms, the aim is to find feasible measures of proven validity and reliability that can be taken from a large sample of animals. The measuring tool or system must be simple, easily operated by trained people and require minimal time and handling of the animal. Non-invasive ratings by human observers to assess a range of animal welfare variables offer some practical advantages as they are inexpensive can be used to integrate multimodal information across time and context and have been shown to be reliable and valid (reviewed by Meagher 2009). However, if observers are required to use complex check lists at each farm visit, the inspection may be too time consuming and discourage producer adoption. Some measures are regarded as less objective than others and have the potential to be affected by the attitudes and experience of the assessor. Thus, whatever measure is chosen, and irrespective of how many observers are required, tools to check for consistency and objectivity between observers are required to ensure robustness.

The ideal welfare assessment system would have the following features:

- Measures underpinned by scientific evidence demonstrating their validity
- A framework for assessing the trade-offs or weightings between different welfare domains
- Integration of the weighted domains to develop an overall welfare index
- Biologically relevant measurement thresholds corresponding to different levels of welfare
- Practical, repeatable and reliable measures that provide an early indication of impending welf concerns.

5.2. On Farm Assessment

A number of different welfare assessment systems have been developed and implemented throughout the world - though none meet the requirements for an ideal system (see above). Several examples are discussed below in order to illustrate some of the methodological issues with animal welfare assessment, and the implications for development of practical welfare assessment protocols. One major observation is that most welfare assurance/assessment schemes focus on how well the farms comply with the given standards and do not make a scientific evaluation of welfare.

Five Freedoms

Traditionally, the Five Freedoms (Farm Animal Welfare Council (FAWC) 1993), or variations of them, have been used as an aspirational framework to guide welfare assessment. The Five Freedoms have been modified to reflect a more pragmatic approach to animal welfare assessment in recent legislation in the UK (Animal Welfare Act, 2006) and include the following needs:

- for a suitable environment (place to live)
- for a suitable diet
- to exhibit normal behaviour patterns
- to be housed with, or apart from, other animals (if applicable)
- to be protected from pain, injury, suffering and disease.

As discussed, the adequacy of this wording for covering the diversity of ethical views is questionable.

Input and Output Measures

Traditionally farm animal welfare assessment has focused on the environmental observation of resources, or inputs provided to the animals on commercial farms. These are indirect measures that typically assess the quality of the management and stockmanship (e.g. space, food) and are attractive because their measurement is quick, simple and reliable. However, direct animal-based measures or outputs which include measures of animals' response to what is provided are thought to more accurately reflect how an animal is coping within its environment. For example, the animals' physical fitness, health or behaviour (e.g. mortality, reproductive performance, and injuries) may give a better indication of the long term adequacy of a production system.

A key advantage of output or animal-based indicators is that variation in conditions within and between geographical regions, production systems and other features can be accommodated with different management practices and yet remain consistent with legislative requirements. For example, variation in climate conditions may result in different demands for the quality of resting area or space allowance. Such differences are difficult to resolve if only defined by resource-based indicators. It might be easier, for instance, to use animal-based indicators to assess the degree of resting comfort under differing management conditions. However, one concern is that is that many animal-based indicators have yet to be demonstrated as valid measures of animal welfare. To date, most of the research on animal-based indicators has been on reliability issues rather than the validity of the measures.

Even the most widely-applied animal-based measures have not been scientifically-validated. Scandinavian and EU legislation provides for the compulsory evaluation of foot health in meat chickens. Foot health is typically achieved by scoring the incidence and severity of foot pad dermatitis (Berg & Algers, 2004), yet until recently there had been no scientific research on the implications of foot (or hock) burn in terms of pain and changes in behaviour. Gait scoring for lameness in meat chickens is another widely-used animal-based indicator of welfare. Some authors (Knowles et al., 2008) have stated that a Gait Score of 3 and above is a sign of poor welfare yet the validity of this measure, too, has only been recently been properly evaluated (Defra, 2012).

Fortunately, progress has been made validating some output measures. The best recent example is the scientific validation of body condition score (as an indicator of chronic hunger) in sheep, (Verbeek et al., 2011, 2012[ab]), dairy cattle (Matthews et al., 2012) and beef cattle (Ferguson et al., 2012). However, there is an ongoing need to continue validating animal-based measures as indicators of animal welfare.

Another important issue in this context is that collecting animal-based indicators typically requires greater effort in comparison with that required for resource-based or management-based indicators, particularly for animal-based indicators collected on-farm. This has been one of the important criticisms of the Welfare Quality® protocol especially amongst practitioners (veterinarians and farmers). One way to make collection of animal-based measures less onerous is to take the measures in the abattoir post-slaughter using automated recording systems (Valros et al., 2004). Although post-slaughter indicators are not able to be used to manage welfare during the animals' life time, the information can be used to

identify issues that can be fed back to the farmer or can be used for risk-oriented control of the farms in question.

The collection of animal-based measures is particularly problematic in extensive grazing systems. Typically, livestock are generally only mustered a few times a year for management and husbandry procedures, with additional inspections occurring with the monitoring of water supplies and fencing. This infrequent monitoring therefore creates major challenges with respect to welfare assessment. Furthermore, the significance of this is further accentuated when considering the profound production challenges that occur in extensive grazing systems such as seasonal variations in food supply, climatic extremes and variability, parasitism and predation (Petherick and Edge 2008). The development of remote animal measurement and monitoring technologies will provide producers with increased capacity to monitor animal movement and possibly health in these extensive environments. Several research studies have shown wireless sensor networks can monitor animal location and health indicators (e.g. rumen temperature) (Mayer et al., 2004) and can estimate behaviour such as landscape avoidance and selection behaviours (Swain et al., 2011).

As stated, most of the welfare monitoring systems that have been developed are based on input or resource measures. They offer practical advantages as it is easier to collect objective observations of resource provision compared with the more subjective assessment of the outcomes (Main & Webster 2011). These resources are presumed to affect animal welfare but links between specific measures of them and the animal welfare status are not clearly understood (Blokhuis et al., 2006). Furthermore, unless there is close correspondence between input variables (e.g. climatic parameters) and output measures (e.g. heat stress), welfare standards based on input measures will not always guarantee good welfare (Main et al 2003; Offner et al 2003).

Stocking density has been a widely-used input parameter in legislation, government-approved welfare codes and in industry, NGO and corporate welfare assurance schemes aimed at protecting animal welfare during intensive-rearing practices. Yet, there has been a paucity of scientifically-credible evidence to determine the appropriateness of stocking density as an indicator of welfare under commercial productions conditions. Its shortcomings were amply demonstrated by Dawkin's and colleagues (Dawkin's et al., 2004; Jones et al., 2005), who examined the effects of stocking density on meat chicken welfare. Contrary to several other studies conducted under laboratory conditions, it was shown that stocking density had no effect on a broad range of welfare measures (e.g. mortality, gait score, podo-dermatitis) at densities used in typical commercial practice. Variation in other input measures (such as environmental temperature and humidity) have been associated with some chicken welfare problems, but as the relationships were weak, they too would not be ideal welfare indicators.

Given there is no agreed gold standard for the determination of welfare for animals, careful interpretation of data collected from a range of parameters is required. The relative weighting assigned to each parameter selected is critical for effective outcomes and comparisons between enterprises. Further discussion on the different approaches to the integration and weighting of parameters into a practical index is required before assessments of welfare taken in the field can be truly effective (part 2 of this review). Once these welfare indices have been scientifically proven, there will be a need to inform retailers and consumers alike so that consumers can make informed decisions on animal products.

Another good example of using a combination of input and output measures is that used in commercial cattle feedlots in Australia to predict and manage heat stress events. By using a

combination of observed local climatic conditions and animal responses to the heat (panting scores), feedlot managers can manage risks and implement strategies to reduce the impact of severe hot weather (Gaughan et al., 2008). Additional indices, namely the heat load index (HLI) and the accumulated heat load (AHL) determine the animal's heat load balance taking into account the duration of daily heat exposure and the availability of natural cooling at night (Gaughan et al., 2008). Using the new HLI and AHL indices that have been incorporated in to a Web-based model, feedlot managers can determine specific heat risk assessments for different cattle genotypes on a daily, pen by pen basis if required. Although the physiological impact of heat stress on beef cattle has been well quantified and provides sufficient evidence to merit intervention, the impact of heat stress on the animal's experience and its affective state remain unknown (Matthews 2008).

The utility of input and output measures is perhaps best summarised by the European Food Safety Authority (EFSA, 2012). The EFSA view is that input measures are more suited to identifying situations that pose a potential hazard to welfare and that output or animal based measures are more appropriate for assessing welfare and evaluating the effects of management procedures to improve welfare. We support this position.

Welfare Quality®

By comparing the assessment systems currently in place with the set of ideal requirements outlined earlier, the limitations of the current measure methodologies are readily apparent. These limitations are becoming more widely acknowledged (e.g. Farm Animal Welfare Council, 2005) and research is underway around the world to address them. The largest research project of this kind in the world is Welfare Quality (Veissier *et al.* 2007) based in the European Union (EU). A unique feature of this project is the linking of an understanding of societal values and concerns about animal welfare in production processes with the development of appropriate measures. Twelve key elements of animal welfare (Table 2) have been identified and these have been shown to encompass all aspects of welfare underlying the value frameworks of a majority of EU citizens (Kjaernes *et al.* 2007). The protocol covers a slightly wider range of animal attributes than the Five Freedoms by including specific categories including good human-animal relationship and a positive emotional state (even though definitions and assessment criteria remain questionable for emotional status). For example, in the Welfare Quality protocol used to assess the welfare of pigs at slaughter (Velarde and Dalmau, 2012), positive emotional state is measured by recording the incidence of 'reluctance to move' or 'turning back' activities. Clearly, these behaviours reflect aversive experiences of the animals and the absence of these behaviours does not necessarily imply a positive emotional state.

Currently, Welfare Quality is focussing more on the practical aspects of measurement (i.e. feasibility, reliability) than on validity. While these are important issues that need addressing, there is also a clear requirement to address the remaining deficiencies in welfare assessment methodologies.

Table 2: Welfare principles, criteria and some examples of potential measures for each welfare criterion

Principle	Welfare criteria	Examples of potential measures
Good feeding	1. Absence of prolonged hunger	Body condition score
	2. Absence of prolonged thirst	Access to water
Good housing	3. Comfort around resting	Frequencies of different lying positions, standing up and lying down behaviour
	4. Thermal comfort	Panting, shivering
	5. Ease of Movement	Slipping or falling
Good health	6. Absence of injuries	Clinical scoring of integument, carcass damage, lameness
	7. Absence of disease	Enteric problems, downgrades at slaughter
	8. Absence of pain induced by management procedures	Evidence of routine mutilations such as tail docking and dehorning, stunning effectiveness at slaughter
Appropriate behaviour	9. Expression of social behaviours	Social licking, aggression
	10. Expression of other behaviours	Play, abnormal behaviour
	11. Good human-animal relationship	Approach and/or avoidance tests
	12. Positive emotional state	Qualitative Behavioural Assessment (QBA)

A number of different welfare assessment systems have been developed and applied on-farm - though none meet the requirements for an ideal system. Traditionally these systems are input-based focusing on assessments of the animals' environment and resources. Whilst useful, these assessments have limitations and there has been increased emphasis on developing and applying animal-based or output measures. Systems based on the integration of input- and output-based measures are starting to be applied – EU Welfare Quality®. Although more comprehensive, on-going demonstration of validity and practicability is required.

On-Farm Welfare Quality Assurance Schemes

Formal welfare assessment and audits of animals on farms may be required for voluntary farm assurance schemes or to ensure relevant welfare legislation and industry standards are enforced. Farm assurance can affect the welfare of animals through encouraging improvements in welfare by setting good standards of provisions for animals and encouraging continuous improvement. If farm produce is to be certified with a label that implies the produce is derived from farms meeting certain

conditions, including welfare, consumers and trading partners must be assured that such farms are monitored and audited against an agreed set of welfare indices. Farm assurance bodies claim to offer whole chain assurance from the farm to the consumer – farm to fork- encompassing farmer, haulers, abattoirs and suppliers (Whay, 2008). Quality Assurance (QA) is defined as 'a planned and systematic set of activities to ensure that requirements are clearly established and the defined process complies with these requirements' (Isixsigma 2012). Farm QA schemes were first developed in the UK in the early 1990's and currently, there are 12 voluntary schemes assuring consumers that the food is of high quality while upholding good animal welfare and environmental standards (Hubbard 2012). The longest standing assurance scheme is the UK based RSPCA's Freedom Foods which covers a range of species.

Farm assessment is an increasingly valuable component of the agri-food industry for creating quality driven food markets (Buller and Roe 2012) yet the success of such markets will rely heavily on the integrity of the assessment and auditing process. These assurance schemes are recognised as the key tool for assessing on-farm welfare (Veissier et al., 2008) and allow consumers 'buying power' as they can make better informed choices when purchasing animal products. The bodies that set the standards have sometimes been linked to large retailers and employed inspectors to check compliance with rules laid out in the standards (Whay 2008). Different QA schemes place different emphasis on food safety, animal welfare and the environment. For example, the RSPCA Freedom Foods is primarily designed to ensure high standards of animal welfare, but the Red Tractor Scheme is designed to ensure compliance with food safety, sustainability and environmental protection, in addition to animal welfare standards. Hence the development of robust monitoring protocols for welfare and husbandry underpins the effectiveness of any welfare–based quality assurance.

In Australia there has been considerable effort towards the development of standards and/or QA programs that incorporate animal welfare, both nationally (e.g. Barnett and Glatz, 2004; Edge et al., 2008) and internationally (e.g. Blokhuis et al., 2003; Main et al., 2003). A major challenge is for QA programs is the interpretation of data within a variety of production systems.

On farm assurance requires an inspector to visit farms to gather evidence in terms of records and health data, to observe management and to assess of a number of individuals on each farm audited as a representative sample. The RSPCA Freedom Foods scheme has laid out standards of animal management, based around the Five Freedoms that if complied with, are believed to inevitably lead to good animal welfare. The Freedom Food scheme covers every stage of a farm animal's life; each stage governed by strict and compulsory RSPCA welfare standards covering handling, transport and slaughter and now exists in Europe and Australasia. Products labelled with Freedom Food logo are available for purchase for customers shopping with an ethical agenda including animal welfare. These products are sold at a higher price and it was initially thought that farmers who joined the schemes would be able to command a premium price, but with the exception of beef products, the premium payment has not yet filtered through to the producer (Whay 2008). Many farmers now perceive farm assurance as a costly time consuming exercise with which they have no choice but to comply (Whay 2008). Freedom Foods and other schemes including those used by McDonald's continue to implement animal-based welfare assessments within its scheme using protocols based on the Bristol Welfare Assurance Program. These schemes include those applied to animals in abattoirs which appear to have been effective in improving welfare (Grandin 2007).

The Austrian Animal Needs Index (ANI) is an example of an on-farm assessment program which has been used in Europe (Bartussek et al., 1999) and Asia (Seo et al., 2007). The current version for cattle is called ANI35L/2000-cattle and has been widely used for certification and legislative purposed

for many years. The ANI system has been in use for dairy and beef cattle, laying hens and pigs. This index has the following five components (sheets) to assess animal welfare:

- sheet 1, affording movement and locomotion (Locomotion),
- sheet 2, affording social interaction (Social interaction),
- sheet 3, type and condition of flooring (Flooring),
- sheet 4, light and air conditions (Light and Air) and
- sheet 5, stockmanship (Stockmanship).

This system assesses the welfare level on farms based on environmental parameters, not the actual state of the animals. These categories are assessed and recorded on each evaluation sheet by the assessor, each visit taking one hour. Points are assigned to several parameters within each of the five categories. The total of the points in all sheets is the ANI score with high ANI scores indicating better welfare levels. Correlations between the ANI score and behaviour and health parameters have provided some validity of the index (Ofner 2003). However, it remains unsuitable for assessing restricted housing such as battery cages for layer hens or sow crates because the scoring system requires minimum standards to be fulfilled. Certain assessment parameters have proven to be difficult to judge especially in the stockmanship category, as they included subjective criteria for scoring items such as cleanliness, floor slipperiness and animal health (Seo et al, 2007). Thus, much more detailed information about evaluating such criteria and the relevance of the criteria to animal well being is needed.

The Bristol Welfare Assurance Programme is a good example of a more comprehensive animal-based assessment scheme (www.vetschool.bris.ac.uk/animalwelfare). The protocols for monitoring farm welfare here are very detailed, species specific and based on the principles of the Five Freedoms. These operational welfare assessment protocols are primarily based on animal parameters that have been developed, initially for dairy cows (Whay et al., 2003). For example, for dairy cattle under the principle of freedom from hunger, a body condition assessment is made. Under the provision of freedom from discomfort, a measure of the number of painful conditions such as swollen hocks, and swollen udders are noted. The methodology of how each indicator should be measured is provided. The information gathered from the comprehensive indices of welfare from each farm assessed is circulated to 50 experts who are asked to indicate the herd prevalence which would indicate a welfare problem. For example, what incidence of lameness or percentage of thin cows on dairy farms would the experts recommend intervention at the herd level? Thus, the interpretation of the significance of any health criteria is paramount.

Quality assurance programs can have a role in influencing animal welfare through the standards they set and by providing incentives for good animal welfare. Although there is no evidence in the UK yet, a farm that shows an unacceptably high cost to animal welfare could lose its certification status (Whay 2008). For a scheme to succeed it must operate both on the farm and at the retail end so any added-value is passed on at every link in the food chain, to reward the farmers by informed consumers. Programs must not only incorporate a means to identify the prevalence of a welfare problem but must ensure effective intervention for continuous improvement. Webster and Main (2011) state that many of the assurance schemes are still young and as yet there is little evidence from which to truly assess their impact.

In addition, farm assurance could act as a route for information and knowledge transfer to the consumer, and back to farmers to engender pride in good achievements. Buller and Roe (2012) describe the increasing trend for animal welfare to be commodified, that is for welfare to be a 'value-added' component. However, few retailers believe that 'welfare sells' and reject the notion of a

standalone 'welfare label'. Increasing support for free range systems for egg production, which is a system- or input-based measure of welfare and one that may not always represent high welfare for the birds, questions the role for the new range of output-based measures in the minds of consumers. Thus, there will be an additional need to convey messages to consumers, perhaps through labelling, regarding the importance of using animal outcomes as measures of welfare.

Welfare auditing introduces an additional layer of independent monitoring. Inspection for farm assurance involves measuring or testing parameters on the day of the visit and then comparing data to a standard. However, welfare auditing involves providing assurance that the practices observed on the day of the inspection are likely to be sustained in the future. This involves ensuring problems are prevented, re-evaluating inspection outcomes and involves herd health planning (Whay 2008). Finally, an effective assessment system involves a means of integrating the weighted input- and output-based components into an overall index and concurrently determining relevant thresholds for at risk animals. Does farm welfare assurance deliver good animal welfare?

Just how effective these assurance schemes are on ensuring high animal welfare standards is unclear. A recent UK report showed an association between farm assurance scheme membership and increased compliance with welfare codes and legislation from 2003-2008 (KilBride et al., 2011). However, there were differences between countries and associations varied across enterprise types and there was insufficient evidence to analyse the effect of schemes standards that exceed welfare codes.

The impact of the RSPCA Freedom Food scheme on the welfare of dairy cattle was studied and outcomes of measures compared between farms belonging to the Freedom Food scheme or other schemes (Main et al., 2003). The Freedom Food farms performed less well for welfare indicators including hock injuries, lameness and restrictions in rising behaviour, but Freedom Food farms performed better in terms of indicators of mastitis, cleanliness and body condition. However, regardless of the scheme, welfare problems remained prevalent indicating that setting standards of provision alone is insufficient to ensure good welfare (Whay 2008). The Scottish Agricultural College (2007) compared the welfare of dairy cows in organic milk production systems and showed that levels of lameness and hock damage were lower on organic farms as a result of shorter winter housing periods and a higher age of first calving heifers which are both elements of the more extensive housing approach described in organic standards. Whay (2008) concludes that with the exception of Freedom Foods and Soil Association Certification there is little evidence of a genuine effort among farm assurance schemes in the UK to use their role to push for animal welfare improvement on farms.

Another important question when assessing the welfare status of livestock on farms is: how many animals should be sampled in the assessment procedure? The answer is not simple as it depends on finding a balance between time constraints in conducting the assessment and the level of accuracy required at the individual or farm level. There have been several recent studies exploring this question. Main et al. (2010) reported that, for larger herds, a sample size of 100 cows is required. However, if the goal is to detect farms, rather than individual animals, with a lameness problem, then focussing on the numbers of cows with severe lameness at the end of milking is an efficient strategy. Alternatively, concentrating assessments on the middle third of the milking order gives an accurate estimate of the prevalence of lameness in a herd. Vasseur et al. (2012) have shown that to estimate lying time in dairy cattle accurately requires four days of continuous (automated) data collection and that parity and stage of lactation need to be taken into account when selecting animals.

Finally, all systems of scoring animals (farm, abattoir, or research) for welfare outcomes are underpinned by value-based decisions. The evaluative nature of scoring animals does not mean they should be rejected but that we are required to make the ethical judgements clear (Veissier et al., 2011). Ethical values influence both the choice of measures to record and their interpretation. A model for the overall measurement of livestock welfare is needed where the assessment will rely on indicators that cover multiple dimensions including health, physical comfort and expression of behaviours etc. Furthermore, the importance, or weighting, of each dimension in the model is inherently a value-based decision. Thus consultations between social scientists and animal scientists are recommended as demonstrated in the Welfare Quality Project (reviewed by Veissier et al., 2011). Value- based decisions are required in determining whether the assessment is made at the individual or farm level, and whether the condition of the average or the worse-off animal(s) is considered. In general, welfare is a concept that applies to the individual as it is regarded as a subjective experience, but when we rate welfare at a farm level we typically mean the welfare of all animals on that enterprise collectively. Thus one option is to make an aggregate score based on information at the individual level, so that the proportion of animals in a good versus bad state is measured. Alternatively, the farm may be evaluated at the criterion level, so that the performance of farms against set criterion (presence of disease, milk yield) can be compared. A further consideration is whether aggregation within a criterion is more important than the range from better off and worse off animals. Decisions must be made as to whether a farm that has a low percentage of animals suffering from a severe condition (disease) is rated above or below one that has all animals suffering equally from a mild welfare problem (poor nutrition).

The emergence of ethical consumerism has been a driving factor behind the development of welfare assurance systems. Such schemes are typically based on independent audits of animal resources or of both resource and animal-based variables. There is still a question over whether such systems genuinely facilitate improved animal welfare.

5.3. Legislative Regulations in Australia

In Australia, animal welfare legislation is state or territory based and is primarily concerned about the protection of animals from cruelty. Under Australian law, the stockperson or person responsible for animals is designated as the person in charge and they have a "duty of care" to the animal(s). Thus, livestock producers have a legal obligation to prevent harm and be aware of the welfare state of animals in their care both in intensive and extensive production systems. In addition, national guidelines exist which underpin the legislation and detail expected practices described in the Model Codes of Practice for the Welfare of Animals and are adopted by each state or territory. The existence of these Codes provides a form of welfare regulation and all Australian states and territories have agreed to work on enforcing national minimum standards for livestock. Although a limitation with these codes is that they are voluntary. A new initiative under the auspices of Australian Animal Welfare Strategy (AAWS 2005) involves the translation of existing codes into legislated national standards. An additional option currently being discussed is to legislate for key animal welfare requirements to be delivered via Quality Assurance (QA) programs or a licence using a co-regulatory approach (Edge et al., 2008). This would imply the government then oversees a guarantee that the legislation delivers the required outcomes. Currently, in most states and territories of Australia producers are not subject to routine government welfare inspections. However, state based assessments made against the animal welfare Codes of Practice (which are primarily resource-based guidelines) have been recently proposed. For example in Tasmania, unannounced animal welfare inspection now occurs on all commercial poultry farms and pig farms with 50 or more pigs, for the purpose of assessing animal welfare. Under this program, farms are subject to inspection about once every 1-2 years to check for compliance with the relevant Code of

Practice for Pigs or Poultry (Department of Primary Industries Parks Water and Environment, Tasmania 2012). The development of a national QA system would reduce the need for multiple state regulatory enforcements of standards and provide an opportunity to consolidate legislative and commercial requirements (Edge 2008).

In Australia, all model codes of practice for farm animal welfare are being replaced with national standards and guidelines. Standards will be enforceable under law and these represent the minimum level of animal welfare required.

6. Conclusions

Animal welfare is a complex human construct. There are three conceptual frameworks that have been applied for assessing farm animal welfare:

- Biological functioning – normality as evidenced through measures of behaviour, physiology, health and productivity
- Affective states – as evidenced through measures of abnormal behaviours, affective states (positive and negative feelings) and cognitive function, and
- Naturalness – as evidence by attributes of the animal, in particular normal behavioural repertoires, and by attributes of its environment

These conceptual frameworks are not mutually exclusive but are complementary in our fundamental understandings of animal welfare.

The viability of livestock farming requires practices that are not only productive, profitable and sustainable but fit with society's expectations on ethical dimensions such as animal welfare. Transparent demonstration of how these expectations have been met will be paramount in the future.

Society comprises a range of stakeholders including: governmental regulatory and policy making bodies, producers, marketers, citizens/consumers, scientists, retailers/service providers, non-Governmental organisations (NGOs) and animal advocacy groups. Given this diversity, it is extremely challenging to develop a common welfare assessment framework that meets the requirements of all stakeholders.

A useful first step for undertaking a welfare assessment is to define the purpose or reason the assessment is being undertaken. Four board categories of reasons for undertaking assessments are:

1. To inform the development and implementation of policies and regulations
2. For assessment and quality assurance of farm animal practices
3. For research purposes
4. To enable better on-going welfare management on farm

A second valuable step in welfare assessment that is not currently practiced would be articulation of the values and ethical frameworks used for selection of assessment criteria and for interpretation of data. This step acknowledges that welfare assessment is an evaluative process in which values influence the choice of measures, their interpretation and their weighting when the measures are combined in any legislative standard, QA assessment system or research methodology. Thus engagement and improvement in the quality of dialogue amongst stakeholders is needed so that the reasons for the choice of assessment measures are more clearly understood.

For on-farm welfare assessment it is evident that there is no one comprehensive, fully-validated system for evaluating the welfare of the diversity of species, production environments and management systems used in livestock enterprises. This does not imply that we are starting from a zero base, as research has shown that assessments combining health and production data, observation of behaviour and physical appearance of animals within a group offer reliable and feasible tools for welfare assessment . The strategic combination of input or resource and output or animal-based measures is important, particularly for welfare risk assessment (e.g. assessments of pasture/forage availability + body condition score in cattle or sheep). In addition to the requirement to demonstrate the validity of these measures or systems, efforts should also be directed to improving the interpretation and applicability of assessment systems within the various livestock enterprises. Priority should be given to exploring avenues to improve the reliability whilst reducing the complexity and invasiveness of methodologies. The development and application of remote automated data capture systems is central here in both extensive and intensive animal production systems.

Preeminent in welfare science is the development of a better understanding of the emotional range and valence in livestock species. To that end, the ongoing development and validation of behavioural and cognitive methodologies is essential. This could be greatly enhanced through the integration of neuroscience disciplines particularly with respect to the validation of these methodologies and the development of novel measures (e.g. lateralisation). The capacity to assess affective states in production environments is a current limitation. However, Qualitative Behavioural Assessment (QBA) would appear to offer most promise to date and further investigation is warranted. The convergence between affective state and biological function, as demonstrated in recent research, represents an important development in welfare science. Indeed, this approach based on the integration of biological function and affective state measures provides a more robust methodology to examine the welfare impacts of a particular production or husbandry issue. Such an approach should be encouraged in future welfare research.

In summary, the development of appropriate welfare assessment methodologies that are credible to all stakeholders will be built on a better understanding of:

- changes in physical health and biological functioning that correspond with different levels of welfare;
- the capacities of livestock to experience negative and positive mental states and associated levels of welfare;
- the ways that separate measures and welfare attributes can be weighted and integrated to give an overall index of welfare;
- how these can be practically implemented in the production environment.

7. References

Aerts S, Lips D, Spencer S, Decuypere E, De Tavenier J (2006). A new framework for the assessment of animal welfare: Integrating existing knowledge from a practical ethics perspective. *Journal of Agricultural and Environmental Ethics* **19,** 67-76.

Appleby MC and Hughes BO (Eds) (2011). 'Animal Welfare.' (CAB International: Wallingford, UK)

Appleby, MC, Hughes, BO and Elson, HA (1992). Poultry Production Systems. CAB International, Wallingford, Oxfordshire, UK.

Barbieri S and Nassuato C (2007). Basic information for the development of the animal welfare risk assessment guidelines [Online]. European Food Safety Authority EFSA/AHAW/2006/01. Available at http:/www.efsa.europa.eu/EFSA/DocumentSet/AHAW_report_basicinfo_AWRA_en,0.pdf (accessed 22 December 2007)

Barnett JL and Glatz PC (2004). Developing and implementing a welfare audit in the Australian chicken meat industry. In Measuring and Auditing Broiler Welfare (C. Weeks & A. Butterworth eds). CAB International, Wallingford, pp 231-240.

Barnett JL, Hemsworth PH, Cronin GM, Jongman EC and Hutson GD (2001). A review of the welfare issues for sows and piglets in relation to housing. *Australian Journal of Agricultural Research* **52,** 1-28.

Barnett, JL (2003). Studying stress to assess animal welfare. In: Paterson, J.E. (ed.) Manipulating Pig Production IX. Australasian Pig Science Association, Werribee, Australia, pp. 107-114 and 116-120.

Barnett, JL and Hemsworth, PH (2009). Welfare monitoring schemes: using research to safeguard welfare of animals on the farm. *Journal of Applied Animal Welfare Science* **12**,114–131.

Bartussek H (1999). A review of the animal needs index (ANI) for the assessment of animals' well-being in the housing systems for Austrian proprietary products and legislation *Livestock Production Science*, **61**, 179-192.

Bartussek H, Leeb CH and Held S (2000). Animal needs index for cattle.

Berridge KC (1996). Food reward: brain substrates of wanting and liking. *Neuroscience and Biobehavioural Reviews* **20,** 1-25.

Berridge KC (2003). Pleasures of the brain. *Brain and Cognition* **52,** 106-128.

Berridge KC, Robinson KC (2003). Parsing reward. *Trends in Neuroscience* **26,** 507-513.

Berry, D.P., Bermingham, M.L., Good, M., More, S.J., 2011. Genetics of animal health and disease in cattle. Irish Veterinary Journal 64.

Blokhuis HJ, Jones RB, Geers R, Miele M and Veissier I (2003). Measuring and monitoring animal welfare: transparency in the food product quality chain. *Animal Welfare* **12**, 445-455.

Blokhuis, HJ, Jones, RB, Veissier, I, and Geers R (2006). Cost Action 846 Measuring and Monitoring farm animal welfare. K.U. Leven.

Boissy A, Manteuffel G, Jensen MB, Oppermann M, Spruijt B, Keeling LJ, Winckler C, Forkman B , Dimitrov I, Langbein J, Bakken M, Veissier I and Aubert A (2007). Assessment of positive emotions in animals to improve their welfare. *Physiology & Behaviour* **92**, 375–397

Botreau R, Bonde M, Butterworth A, Perny P, Bracke MBM and Capdeville J (2007). Aggregation of measures to produce an overall assessment of animal welfare. Part 1: a review of existing methods. *Animal* **1,** 1179-1187.

Bracke MBM, Spruijt BM, Metz JHM and Schouten WGP (2002). Decision support system for overall welfare assessment in pregnant sows A: Model structure and weighting procedure. *Journal of Animal Science* **80,** 1819-1834.

Brambell Committee 1965. Report of the etchnical committee to enquire into the welfare of animals kept under intensive livestock husbandry systems. Her Majesty's Staionary Office, London.

Broom, DM (1986). Indicators of poor welfare. *British Veterinary Journal* **142**, 524-526.

Broom, DM (1996). Animal welfare defined in terms of attempts to cope with the environment. *Acta Agriculturae Scandinavica, Section A - Animal Science* **27**, 22-28.

Burman O, McGowan R, Mendl M, Norling Y, Paul E, Rehn T and Keeling L (2011). Using judgement bias to measure positive emotional states in dogs. *Applied Animal Behaviour Science* **132,** 160-168.Cabanac, M (1979). Sensory pleasure. *Quarterly Review of Biology* **54**, 1-29.

Burrow, H.M., 2012. Importance of adaptation and genotype X environment interactions in tropical beef breeding systems. Animal 6, 729-740.

Cacioppo JT, Klein DJ, Berston G and Hatfield E (1993). The psychophysiology of emotion. In: Lewis, M. and Haviland, JM (eds.) Handbook of Emotions. Guildford Press, New York, USA, pp. 119-142.

Canali E and Keeling L (2009). Welfare Quality® project: from scientific research to on farm assessment of animal welfare. *Italian Journal of Animal Science* **8**, 900-903

Champion RA and Matthews LR (2007). An operant-conditioning technique for the automatic measurement of feeding motivation in cows. *Computers and Electronics in Agriculture* **57**, 115–122.

Clarke N, Barnett C, Cloke P and Malpass A (2007). Globalising the consumer: Doing politics in an ethical register. *Political Geography* **26**, 231-249.

Colditz IG, Paull DR, Lloyd JB and Fisher AD (2010). Physiological and behavioural effects of intradermal injection of sodium lauryl sulfate as an alternative to mulesing in lambs. *Australian Veterinary Journal* **88**, 483–489.

Colditz, IG and Le Jambre, LF (2008). Development of a faecal occult blood test to determine the severity of *Haemonchus contortus* infections in sheep. *Veterinary Parasitology* **153**, 93-99.

Cooper J J and Albentosa MJ. (2003). Behavioural priorities of laying hens. *Avian Poultry Biology Review* **14**,127-149.

Croney, C.C., Anthony, R., 2010. Engaging science in a climate of values: Tools for animal scientists tasked with addressing ethical problems. Journal of Animal Science 88, E75-E81.

Croney CC, Apley M, Capper JL, Mench, JA and Priest, S (2012). BIOETHICS SYMPOSIUM: The ethical food movement: What does it mean for the role of science and scientists in current debates about animal agriculture? Journal of Animal Science 90, 1570-1582.

Dairy Insight (2007). Dairy industry strategy for animal welfare. Dairy Insight, Wellington, NZ.

Danbury TC, Weeks CA, Chambers JP, Waterman-Pearson AE and Kestin SC (2000). Self-selection for the analgesic drug carprofen by lame broiler chickens. *Veterinary Record* **146,** 307-311.

Dantzer R and Mormède P (1983). Stress in farm animals: a need for reevaluation. *Journal of Animal Science* **57**, 6-18.

Dawkins M (1980). Animal Suffering: The Science of Animal Welfare. Chapman and Hall, New York, USA.

Dawkins M (1983). Battery hens name their price: consumer demand theory and the measurement of animal needs. *Animal Behaviour* **31**, 1195-1205.

Dawkins M S (1990). From an animal's point of view: Motivation, fitness, and animal welfare. *Behavioural Brain Science* **13**, 1–9.

Dawkins MS (2003). Behaviour as a tool in the assessment of animal welfare. *Zoology* **106**, 383-387.

Dawkins MS (2006). Through animal eyes: What behaviour tells us. *Applied Animal Behaviour Science* ***100****, 4-10.*

Dawkins MS, Donnelly CA and Jones TA (2004). Chicken welfare is influenced more by housing conditions than by stocking density. *Nature* **427**, 342-344.

Dawkins, M (1977) Do hens suffer in battery cages? Environmental preferences and welfare. Animal Behaviour 25, 1034-1046.

DEFRA (2010). Foot pad dermatitis and hock burn in broilers: risk factors, aetiology and welfare consequences. Report AW1137.

Denton DA, McKinley MJ, Farrell M, Egan GF. The role of primordial emotions in the evolutionary origin of consciousness. Consciousness and Cognition 18, 500–14, 2009

Denton, D. A. (2006). The primordial emotions: The dawning of consciousness. Oxford University Press.

Department of Primary Industries Parks Water and Environment, Tasmania 2012 Animal Welfare Inspections of Intensive Farms http://www.dpiw.tas.gov.au/inter.nsf/WebPages/LBUN-8SM6TA?open.

Dominik, S., Henshall, J.M., Hayes, B.J., 2012. A single nucleotide polymorphism on chromosome 10 is highly predictive for the polled phenotype in Australian Merino sheep. Animal Genetics 43, 468-470.

Douglas C, Bateson M, Walsh C, Bédué A and Edwards SA (2012). Environmental enrichment induces optimistic cognitive biases in pigs. *Applied Animal Behaviour Science* **139**, 65–73.

Doyle RE, Fisher AD, Hinch GN, Boissy A and Lee C (2010). Release from restraint generates a positive judgement bias in sheep. *Applied Animal Behaviour Science* **122**, 28–34.

Doyle RE, Hinch GN, Fisher AD, Boissy A, Henshall JM, Lee C (2011[a]). Administration of serotonin inhibitor p-Chlorophenylalanine induces pessimistic-like judgement bias in sheep. *Psychoneuroendocrinology* **36**, 279–288.

Doyle RE, Lee C, Deiss V, Fisher AD, Hinch GN and Boissy A (2011[b]). Measuring judgement bias and emotional reactivity in sheep following long-term exposure to unpredictable and aversive events. *Physiology and Behaviour* **102**, 503-510.

Duncan I J H (2005). Science-based assessment of animal welfare: farm animals. *Rev. Sci. Tech. Off. Int. Epiz.* **24**, 483-492.

Duncan IJ (1998). Behavior and behavioral needs. *Poultry Science* **77**, 1766–1772

Duncan IJH (1978). The interpretation of preference tests in animal behaviour. *Applied Animal Ethology* **4**,197-200.

Duncan IJH (2004). A concept of welfare based on feelings. In: Benson, G.J. and Rollin, B.E. (eds.) The Well-Being of Farm Animals: Challenges and Solutions. Blackwell Publishing, Ames, Iowa, USA, pp 95-101.

Duncan IJH (2006). The changing concept of animal sentience. *Applied Animal Behaviour Science* **100**, *11-19*.

Duncan IJH (2006). the changing concept of animal sentience. *Applied Animal Behaviour Science* **100**, 11-19.

Duncan IJH and Fraser D (1997). Understanding Animal Welfare. In: Appleby, MC and Hughes, BO (eds.) Animal Welfare. CAB International, Wallingford, Oxfordshire, UK, pp. 19-31.

Edge MK, Hemsworth PH, and Barnett JL (2008). Verifying legislative and customer requirements utilising animal welfare quality assurance. *Australian Journal of Experimental Agriculture* **48,** 1022-1027.

Edwards SA (2007). Experimental welfare assessment and on-farm application. *Animal Welfare* **16**, 111-115

EFSA (European Food Safety Authority) (2005) Welfare Aspects of Various Systems for Keeping Laying Hens. Annex to the EFSA Journal 197, 1-23. EFSA-Q-2003-92.

EFSA Panel on Animal Health and Welfare, 2012. Statement on the use of animal-based measrues to assess the welfare of animals. EFSA Journal 10, 2767-2796.

Eurobarometer (2005) Attitudes of consumers towards the welfare of farmed animals [Online]. Available at: http://ec.europa.eu/food/animal/welfare/euro_barometer25_en.pdf) (accessed 30 August 2007)

Evans A and Miele M (Eds) (2007). Consumers' views about farm animal welfare. Part I: National reports on focus group research. Welfare Quality® Reports No. 4, Cardiff.

Fahmy, M.H., Bernard, C., 1971. Causes of mortality in Yorkshire pigs from birth to 20 weeks of age. Can. J. Anim. Sci. 51, 351-359.

Farm Animal Welfare Council (FAWC) (1993). Second Report on priorities for animal welfare research and development. Department for Environment, Food and Rural Affairs Publications, London.

Farm Animal Welfare Council (FAWC) (2009) Farm animal welfare in Great Britain; Past , present and future. Farm Animal Welfare Council, London.

Farm Animal Welfare Council (FAWC) (2005). Report on the welfare implications of welfare assurance schemes. Farm Animal Welfare Council, London.

Febrer K, Jones TA, Donnelly CA and Dawkins MS (2006). Forced to crowd or choosing to cluster? Spatial distribution indicates social attraction in broiler chickens. *Animal Behaviour* **72,** 1291-1300.

Fisher AD, Stewart M, Verkerk GA, Morrow CJ and Matthews LR (2003). The effects of surface type on lying behaviour and stress responses of dairy cows during weather-induced removal from pasture. *Applied Animal Behaviour Science* **81,** 1-11.

Fisher AD, Verkerk GA, Morrow CJ and Matthews LR (2001). The effects of feed restriction and lying deprivation on pituitary-adrenal axis regulation in lactating cows. *Livestock Production Science* **73,** 255-263.

Fraser D (1995). Science, values and animal welfare: Exploring the "inextricable connection". *Animal Welfare* **4,** 103-117.

Fraser D (2003[a]). Emerging animal welfare standards and their implications for animal hygiene. XI ISAH Congress in Animal Health, 23 - 27 February 2003, Mexico City, pp 61-66.

Fraser D (2003[b]). Assessing animal welfare at the farm and group level: the interplay of science and values. *Animal Welfare* **12**, 433-443.

Fraser D (2008). Understanding Animal Welfare: The Science in its Cultural Context. Wiley-Blackewell, Chichester, West Sussex, UK.

Fraser D and Duncan I J H (1998). 'Pleasures', 'pains' and animal welfare: toward a natural history of affect. *Animal Welfare* 7, 383-396.

Fraser D and Matthews LR (1997). Preference and motivation testing. In 'Animal Welfare' (Eds Appleby MC, Hughes BO) pp. 159–173. (CAB International, Wallingford, UK)

Fraser D and Matthews LR (1997). Preference and motivation testing. In: Appleby, MC and Hughes, BO (eds.) Animal Welfare. CAB International, Wallingford, Oxforshire, UK, pp. 159-174.

Fraser D, Weary DM, Pajor EA and Milligan BN (1997). A scientific conception of animal welfare that reflects ethical concerns. *Animal Welfare* **6,** 187-205.

Gaughan JB, Mader TL, Holt SM and Lisle A (2008). A new heat load index for feedlot cattle. *Journal of Animal Science* **86**, 226-234

Gornati R, Papis E, Rimoldi S, Chini V, Terova G, Prati M, Saroglia M and Bernardini G (2005). Molecular markers for animal biotechnology: sea bass (*Dicentrarchus labrax*, L.) HMG-CoA reductase mRNA. *Gene* **344**, 299-305.

Grandin T (2007). Introduction: Effect of customer requirements, international standards and marketing structure on the handling and transport of livestock and poultry. In 'Livestock handling and transport 3rd Edition' (Ed. T Grandin) pp.1-18. CAB International:Wallingford, UK)

Greiner M and Gardner IA (2000). Epidemiologic issues in the validation of veterinary diagnostic tests. *Preventive Veterinary Medicine* **45**, 3-22.

Guatteo, R., Levionnois, O., Fournier, D., Guemene, D., Latouche, K., Leterrier, C., Mormede, P., Prunier, A., Serviere, J., Terlouw, C., Le Neindre, P., 2012. Minimising pain in farm animals: the 3S approach - 'Suppress, Substitute, Soothe'. Animal 6, 1261-1274.

Hatcher, S., Eppleston, J., Graham, R.P., McDonald, J., Schlunke, S., Watt, B., Thornberry, K.J., 2008. Higher weaning weight improves post weaning growth and survival in young Merino sheep. Australian Journal of Experimental Agriculture 48, 966-973.

Hemsworth PH and Coleman GJ (2011). Human-Livestock Interactions: The Stockperson and the Productivity and Welfare of Farmed Animals, 2nd Edition. CAB International, Oxford.

Hemsworth PH, Barnett JL, Karlen GMA, Fisher AD, Butler KL and Arnold NA (2009). Effects of mulesing and alternative procedures to mulesing on the behaviour and physiology of lambs. *Applied Animal Behaviour Science* **117**: 20-27.

Hubbard C (2012). Do farm assurance schemes make a difference to welfare? *Veterinary Record* **170**, 150-151.

Hughes BO and Black AJ (1973). The preference of domestic hens for different types of battery cage floor. *British Journal of Poultry Science* **14**, 615-619.

Hursh SR and Winger G (1995). Normalised demand for drugs and other reinforcers. *Journal of the Experimental Analysis of Behavior* **64,** 373-384.

Hutson G D (1981). Sheep movement on slated floors. *Australian Journal of Experimental Agriculture and Animal Husbandry* **21**, 474-479.

Hutson GD (1984). Animal welfare and consumer demand theory: are preference tests a luxury we can't afford? *Animal Behaviour* **32**, 1260-1261.

Huzzy JM, Veira DM, Weary DM and von Keyserlingk MAG (2007). Prepartum behaviour and dry matter intake identify dairy cows at risk of metritis. *Journal of Dairy Science* **90,** 3349-3354.

Isixsigma (2012). Definition of quality assurance. Available at: www.isixsigma.com/dictionary [Verified 19 August 2012]

James, P.J., 2006. Genetic alternatives to mulesing and tail docking in sheep: a review. Australian Journal of Experimental Agriculture 46, 1-18.

Janhunen, M., Kause, A., Jarvisalo, O., 2012. Costs of being extreme - Do body size deviations from population or sire means decrease vitality in rainbow trout? Aquaculture 370, 123-129.

Jarvis S, Moinard C, Robson SK, Baxter E, Ormandy E, Douglas AJ, Seckl JR, Russell JA and Lawrence AB (2006). Programming the offspring of the pig by prenatal social stress: Neuroendocrine activity and behaviour. *Hormones and Behaviour* **49,** 68-80.

Jensen, P., 2010. Domestication, selection, behaviour and welfare of animals - genetic mechanisms for rapid responses. Animal Welfare 19, 7-9.

Jones T A, Donnelly C A and Dawkins M S (2005). Environmental and management factors affecting the welfare of chickens on commercial farms in the United Kingdom and Denmark stocked at five densities. *Poultry Science* **84**, 1–11.

Kaltas GA and Chrousos GP (2007). The neuroendorinology of stress. In: Cacioppo, JT, Tassinary, LG and Berntson, GG (eds.) Handbook of Psychophysiology. Cambridge University Press, Cambridge, UK, pp. 303-318.

Keeling L and Jensen P (2005). Abnormal behaviour, stress and welfare. In The Ethology of Domestic Animals, 2nd ed., Jensen, P (Ed.) CABI Publishing: Wallingford, UK, pp. 85–101.

Kestin SC, Knowles TG, Tinch AE and Gregory NG (1992). Prevalence of leg weakness in broiler chickens and its relationship with genotype. *Veterinary Record* **131,** 190-194.

KilBride AL, Mason SA, Honeyman PC, Pritchard DG, Hepple S and Green LE (2011). Associations between membership of farm assurance and organic certification schemes and compliance with animal welfare legislation. *The Veterinary Record* **170**:152.

Kirkden RD and Pajor EA (2006[a]). Using preference, motivation and aversion tests to ask scientific questions about animals' feelings. *Applied Animal Behaviour Science* **100,** 29-47.

Kirkden RD and Pajor EA (2006[b]). Motivation for group housing in gestating sows. *Animal Welfare* **15,**119-130.

Kirkden RD, Edwards JSS and Broom DM (2003). A theoretical comparison of the consumer surplus and the elasticities of demand as measures of motivational strength. *Animal Behaviour* **65,** 157-178.

Kjaernes U, Miele M, Roex J (Eds.) (2007) Attitudes of consumers, retailers and producers to farm animal welfare. Welfare Quality® Reports No. 2, Cardiff.

Knierim U and Winckler C (2009). On-farm welfare assessment in cattle: Validity, reliability and feasibility issues and future perspectives with special regard to the Welfare Quality approach. *Animal Welfare* **18**:451–458.

Koeck, A., Miglior, F., Kelton, D.F., Schenkel, F.S., 2012. SHORT COMMUNICATION: Genetic association of body condition score with disease resistance in first lactation Canadian Holsteins. Can. J. Anim. Sci. 92, 285-289.

Korte Sm, Oliver B and Koolhaus JM (2007). A new animal welfare concept based on allostasis. *Physiology and Behavior* **92**, 422-428.

Lawrence, AB and Illius AW (1997). Measuring preferences and the problems of identifying proximate needs, Animal Choices. *Animal Science* **20**, 19-26.

Laywel (2006). Overall strengths and weaknesses of each defined housing system for laying hens, and detailing the overall welfare impact of each housing system [Online]. Available at http://www.laywel.eu/web/pdf/deliverable%2071%20welfare%20assessment-2.pdf accessed 4 (Accessed 5 April 2008)

Lea SG (1978). The psychology and economics of demand. *Psychological Bulletin* **85,** 441–466.

LeDoux JE (1996). The Emotional Brain. Simon & Schuster, New York, USA.

Lee C and AD Fisher (2007). Welfare consequences of mulesing of sheep. *Australian Veterinary Journal* **85**, 89-93

Lepherd ML, Canfield PJ, Hunt GB and Bosward, KL, (2009). Haematological, biochemical and selected acute phase protein reference intervals for weaned female Merino lambs. *Australian Veterinary Journal* **87**, 5-11.

Madden GJ, Smethells JR, Ewan EE and Hursh SR (2007). Tests of behavioural economic assessments of relative reinforcing efficacy: economic substitutes. *Journal of the Experimental Analysis of Behavior* **87,** 219-240.

Main DC and J Webster (2011). Assessment, Implementation and Promotion of Farm Animal Welfare In: Management and Welfare of Farm Animals Ed. 5, UFAW Farm Handbook. Wiley Blackwell, Chichester, UK. pp. 564-584

Main DCJ and Mullan S (2011). AssureWel: incorporation of welfare outcomes into UK laying hen assurance schemes. Proceedings of the 5th International Conference on the Assessment of Animal Welfare at Farm and Group Level (Ed.s T Widowski, P Lawliss, K Shephard) Guelph August 2011 pp. 5.

Main DCJ, Barker ZE, Leach KA, Bell NJ, Whay HR and Browne WJ (2010). Sampling strategies for monitoring lameness in dairy cattle. *Journal of Dairy Science* **93**, 1970–1978

Main DCJ, Whay H R, Green LE and Webster AJF (2003). Effect of the RSPCA Freedom Food Scheme on the welfare of dairy cattle. *Veterinary Record* **153,** 227-231.

Main DCJ, Whay HR, Leeb C and Webster AJF (2007). Formal animal-based welfare assessment in UK certification schemes. *Animal Welfare* **16,** 233-236.

Manning L, Chadd SA and Baines RN (2007). Key health indicators for broiler production. *World's Poultry Science Journal* **63,** 46-62.

Mariasegaram, M., Harrison, B.E., Bolton, J.A., Tier, B., Henshall, J.M., Barendse, W., Prayaga, K.C., 2012. Fine-mapping the POLL locus in Brahman cattle yields the diagnostic marker CSAFG29. Animal Genetics 43, 683-688.

Mason GJ, Cooper J and Clarebrough C (2001). Frustrations of fur-farmed mink. Nature 410, 35-36.

Mather, J.A., 2011. Philosophical Background of Attitudes toward and Treatment of Invertebrates. Ilar Journal 52, 205-212.

Mathews A, Mogg K, Kentish J and Eysenck M (1995). Effect of psychological treatment on cognitive bias in generalized anxiety disorder. Behaviour Research and Therapy 33, 293–303

Matthews LR (2007). International drivers for animal welfare change and implications for New Zealand. Proceedings of the New Zealand Society of Animal Production 67, 226-231.

Matthews LR (2008). Methodologies by which to study and evaluate welfare issues facing livestock systems of production. *Australian Journal of Experimental Agriculture* **48**, 1014-1021

Matthews LR and Ladewig J (1994). Environmental requirements of pigs measured by demand functions. *Animal Behaviour* **47**, 713–719.

Matthews LR, Champion R and Davison D (2006). A novel method for automatically quantifying resource requirements of pastoral livestock in naturalistic settings. Proceedings of the 40th International Congress of the International Society for Applied Ethology, Bristol UK, pp. 104.

Matthews LR, Loveridge A and Guerin B (1994). Animal welfare issues and attitudes in New Zealand. Animal Behaviour and Welfare Research Centre, AgResearch, Hamilton New Zealand.

Meagher RK (2009). Observer ratings: Validity and value as a tool for animal welfare research. *Applied Animal Behaviour Science* **119**, 1-14

Meat and Livestock Australia (2007). Heat load in feedlot cattle [Online]. Available at http://www.mla.com.au/NR/rdonlyres/7C62BAA6-45A7-4F8F-8F54-0697F9B80F96/0/TipstoolsHeatloadinfeedlotcattle2006.pdf (Accessed 5 December 2007)

Mellor D J (2012). Animal emotions, behaviour and the promotion of positive welfare states. *New Zealand Veterinary Journal* **60**, 1-8.

Mellor DJ, Cook CJ and Stafford KJ (2000). Quantifying some responses to pain as a stressor. In: Mench, M and Moberg, GO (eds.) Biology of Animal Stress. CAB International, Wallingford, Oxfordshire, UK, pp. 171-198.

Mellor DJ, Patterson-Kane E and Stafford KJ (2009). The Sciences of Animal Welfare, Wiley-Blackwell, Oxford, UK.

Mendl M, Burman OHP, Parker RMA and Paul ES (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science* **118,** 161–18.

Metz, JHM (1983). Food Competition in cattle. In: S.H. Baxter, M.R. Baxter and J.A.C. MacCormack (eds.) Farm Animal Housing and Welfare. Martinus Nijhoff Publ., Dordrecht, The Netherlands, 164-170

Mills DM (2010). Abnormal/abnormality. In: The Encyclopaedia of Applied Animal Behaviour and Welfare, 1st ed. Mills, DM, Marchant-Forde, JN, Morton, DB, Phillips, CJC, McGreevy, PD, Nicol, CJ, Sandoe, P, Swaisgood, RR (Eds.), CABI Publishing: Wallingford, UK, pp. 2–3.

Minero M, Tosi MV, Canali E and Wemelsfelder F. (2009). Quantitative and qualitative assessment of the response of foals to the presence of an unfamiliar human. *Applied Animal Behaviour Science*. **116,** 74-81.

Moberg GP (2000). Biological response to stress: Implications for animal welfare. In 'The biology of animal stress; basic principles and applications for animal welfare'. (Ed GP Moberg, JA Mench) pp1-21. (CABI Publishing: Wallingford, UK)

Moberg GP and Mench JA (2000). (Eds.) 'The biology of animal stress; basic principles and applications for animal welfare'. (CABI Publishing: Wallingford, UK)

Munksgaard L, Jensen MB, Pedersen LJ, Hansen SW and Matthews LR (2005). Quantifying behavioural priorities – effects of time constraints on behaviour of dairy cows, *Bos Taurus*. *Applied Animal Behaviour Science* **92,** 3-14.

Napolitano F, De Rosa G, Braghieri A, Grasso F, Bordi A, and Wemelsfelder F, (2008). The qualitative assessment of responsiveness to environmental challenge in horses and ponies. *Applied Animal Behaviour Science*. **109,** 342-356.

Napolitano F, Knierim U, Grasso F, and G De Rosa G (2009). Positive indicators of cattle welfare and their applicability to on-farm protocols *Italian Journal of Animal Science* **8,** 355-365

Napolitano F, De Rosa G, Braghieri A, Grasso F, Wemelsfelder F (2012). Qualitative behavioural assessment of dairy buffaloes *Applied Animal Behaviour Science* **141**, 91-100

National Animal Welfare Advisory Committee (2007). Animal Welfare (Deer) Code of Welfare. Ministry of Agriculture and Forestry, Wellington.

Nicol CJ, Caplen G, Edgar J and Browne WJ (2009). Associations between welfare indicators and environmental choice in laying hens. *Animal Behaviour* **78**, 413–424.

Nicol CJ, Caplen G, Statham P and Browne WJ (2011). Decisions about foraging and risk trade-offs in chickens are associated with individual somatic response profiles. *Animal Behaviour* 82, 255-262

Nielsen, H.M., Olesen, I., Navrud, S., Kolstad, K., Amer, P., 2011. How to Consider the Value of Farm Animals in Breeding Goals. A Review of Current Status and Future Challenges. Journal of Agricultural & Environmental Ethics 24, 309-330.

OECD/FAO (2012). OECD-FAO Agricultural Outlook 2012-2021. OECD Publishing and FAO.

Ofner E, Amon T, Lins M and Amon B (2003). Correlations between the results of animal welfare assessments by the TGI 35L Austrian Animal Needs Index and health and behavioural parameters of cattle. *Animal Welfare* **12,** 571-578.

OIE, World Animal Health Organization (2008). Terrestrial Animal Health Code, 17th edn. Oxford University Press, New York, USA.

Panksepp J (1998). Affective Neuroscience. The Foundation of Human and Animal Emotions. Oxford University Press, London, UK.

Panksepp J (2005). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness and Cognition* **14**, 30–80.

Petherick C and Edge M (2008). Measuring the welfare of livestock in extensive production systems: can we, should we? AAWS International Animal Welfare Conference Gold Coast

Petherick JC (2006). Animal welfare provision for land-based livestock industries in Australia. *Australian Veterinary Journal* **84**, 379-383.

Phillips PA, Fraser D and Thompson BK (1991). Preference by sows for a partially enclosed farrowing crate. *Applied Animal Behaviour Science* **32**, 35-43.

Phillips PA, Fraser D and Thompson BK (1996). Sow preference for types of flooring in farrowing crates. *Canadian Journal of Animal Science* **76**, 485-489.

Phillips PA, Thompson BK and Fraser D (1988). Preference tests of ramp designs for young pigs. *Canadian Journal of Animal Science* **68**, 41-48.

Phythian CJ, Cripps PJ, Michalopoulou E, Jones PH, Grove-White D, Clarkson MJ, Winter AC, Stubbings LA and Duncan JS (2012). Reliability of indicators of sheep welfare assessed by a group observation method. *The Veterinary Journal* **193**, 257–263.

Plesch G, Broerkens N, Laister S, Winckler C and Knierim U (2010). Reliability and feasibility of selected measures concerning resting behaviour for the on-farm welfare assessment in dairy cows. *Applied Animal Behaviour Science* **126**, 19–26

Pollard JC, Littlejohn RP and Suttie JM (1994). Responses of red deer to restraint in a Y maze preference test. *Applied Animal Behaviour Science* **39**, 63-71.

Price EO (2003). Animal Domestication and Behavior, 1st ed., CABI Publishing: Wallingford, UK, p. 180.

Rauw, W.M., Kanis, E., Noordhuizen-Stassen, E.N., Grommers, F.J., 1998. Undesirable side effects of selection for high production efficiency in farm animals, a review. Livestock Production Science 56, 15-33.

Rousing T and Wemelsfelder F (2006). Qualitative assessment of social behaviour of dairy cows housed in loose housing systems. *Applied Animal Behaviour Science* **101**, 40-53.

RSPCA (2007). Freedom Food [Online]. Available at: http://www.rspca.org.uk/servlet/Satellite?pagename=RSPCA/FreedomFood/FreedomFoodHome page (accessed 22 December 2007)

Rushen J (1986). Aversion of sheep to handling treatments: Paired-choice studies. *Applied Animal Behaviour Science* **16**, 363-370.

Sandøe P, Christiansen SB and Appleby MC (2003). Farm animal welfare: the interaction of ethical questions and animalw elfare science. *Animal Welfare* **12**,469-478.

Sandøe P, Forkman F and Christiansen SB (2004). Scientific uncertainty – how should it be handled in relation to scientific advice regarding animal welfare issues? *Animal Welfare* **13**, S121-S126.

Scott EM, Fitzpatrick JL, Nolan AM, Reid J and Wiseman ML (2003). Evaluation of welfare state based on interpretation of multiple indices. *Animal Welfare* **12,** 457-468.

Scottish Agricultural College (2007). The welfare of dairy cows in organic milk production systems. Final Report Edinburgh Scottish Agricultural College.

Seo T, Date K, Daigo T, Kashiwamura F and Sato S (2007). Welfare assessment on Japanese dairy farms using the Animal Needs Index. *Animal Welfare* **16**, 221-223.

Spoolder H, Rosa G de, Horning B, Waiblinger S and Wemelsfelder F (2003). Integrating parameters to assess on-farm welfare. *Animal Welfare* **12,** 529-534.

Spoolder HA, Burbidge JA, Lawrence AB, Simmins PH and Edwards SA (1996). Individual behavioural differences in pigs: intra- and inter-test consistency. *Applied Animal Behaviour Science* **49,** 185-198.

Stevens B, Barnett JL, Tilbrook A and Hemsworth PH (2009). Effects of deprivation of a preferred resource (feed or social contact) on the biological functioning of pigs. In: Manipulating Pig Production XII. Australasian Pig Science Association, van Barneveld, RJ (eds.), Werribee Victoria, Australia, pp. 28.

Stewart M, Webster JR, Verkerk GA, Colyn JJ and Schaefer AL (2005). Infrared thermography as a non-invasive measure of stress in dairy cows. *Journal of Animal Science* **83**, 374.

Stockman CA, Collins T, Barnes AL, Miller D, Wickham SL, Beatty, DT, Blache D, Wemelsfelder F and Fleming PA (2011). A Qualitative behavioural assessment and quantitative physiological measurement of cattle naïve and habituated to road transport. *Animal Production Science* **51**, 240–249.

Stokes JE, Leach KA, Main DC and Whay JHR (2012). An investigation into the use of infrared thermography (IRT) as a rapid diagnostic tool for foot lesions in dairy cattle. *The Veterinary Journal* (in press).

Swain DL, Friend MA, Bishop-Hurley GJ, Handcock RN and Wark T (2011). Tracking livestock using global positioning systems – are we still lost? *Animal Production Science* **51**, 167-175.

Tannenbaum J (1991). Ethics and Animal-Welfare - the Inextricable Connection. *Journal of the American Veterinary Medical Association* **198**, 1360-1376.

Tannenbaum, J. (2001). The paradigm shift toward animal happiness: what it is, why it is happening, and what it portends for medical research. In: Why Animal Experimentation Matters: The Use of Animals in Medical Research, Paul, E. F. and Paul J (eds.), Transaction Publishers, Somerset, N.J.

Tixier-Boichard, M., Leenstra, F., Flock, D.K., Hocking, P.M., Weigend, S., 2012. A century of poultry genetics. Worlds Poultry Science Journal 68, 307-321.

Turner SP and Dwyer CM (2007). Welfare assessment in extensive animal production systems: challenges and opportunities. *Animal Welfare* **16,** 189-192.

Valros A, Ahlstrom S, Rintala H, Hakkinen T and Saloniemi H (2004). The prevalence of tail damage in slaughter pigs in Finland and associations to carcass condemnations. *Acta Agriculturae Scandinavica, Section A* **54**, 213-219.

van der Harst JE, Baars AM and Spruijt BM. (2003). Standard housed rats are more sensitive to rewards than enriched housed rats as reflected by their anticipatory behaviour. *Behavioural Brain Research* **142,** 151-156.

Vasseur E, Rushen J, Haley DB and de Passillé AM (2012). Sampling cows to assess lying time for on-farm animal welfare assessment. *Journal of Dairy Science* **95**, 4968–4977.

Veissier I, Butterworth A, Bock B and Roe E (2008). European approaches to ensure good animal welfare. *Applied Animal Behaviour Science* **113**, 279-297.

Veissier I, Forkman B, Jones B (Eds.) (2007). Assuring animal welfare; from societal concerns to implementation. Second Welfare Quality® Stakeholder Conference. Welfare Quality®, Netherlands.

Veissier, I., Jensen, K.K., Botreau, R., Sandoe, P., 2011. Highlighting ethical decisions underlying the scoring of animal welfare in the Welfare Quality (R) scheme. Animal Welfare 20, 89-101.

Veissier I, Jensen KK, Batreau R and Sandoe P (2011). Highlighting ethical decisions underlying the scoring of animal welfare in the Welfare Quality scheme *Animal Welfare* **20,** 89-101

Velarde A and Dalmau A (2012).Animal welfare assessment at slaughter in Europe: Moving from inputs to outputs. *Meat Science* **92**, 244–251.

Verbeek E, Oliver MH, Waas JR, McLeay LM, Blache D and Matthews, LR (2012[a]). Reduced cortisol and metabolic responses of thin ewes to an acute cold challenge in mid-pregnancy: Implications for animal physiology and welfare. *PLoS ONE* 7(5): e37315. doi:10.1371/journal.pone.0037315

Verbeek E, Waas J, McLeay L and Matthews LR (2011). Measurement of feeding motivation in sheep and the effects of food restriction. *Applied Animal Behaviour Science* **132**, 121-130.

Verbeek E, Waas JR, Oliver MH, McLeay LM, Ferguson DM and Matthews LR (2012[b]). Motivation to obtain a food reward of pregnant ewes in negative energy balance: Behavioural, metabolic and endocrine considerations. *Hormones and Behavior* doi.org/10.1016/j.yhbeh.2012.06.006.

Verkerk G, Tucker C, Kendall P, Webster J, Bloomberg M, Rogers A, Stewart M, Davison D and Matthews L (2006). Physical environments and dairy cow welfare - Lessons from New Zealand Research. *Proceedings of the Society of Dairy Cattle Veterinarians of the New Zealand Veterinary Association* **23,** 35-46.

von Borell E, Broom DM, Csermely D, Dijkhuizen AA, Edwards SA, Jensen P, Madec F and Stamataris C (1997). The welfare of intensively kept pigs. Report of the Scientific Veterinary Committee. (European Union: Brussels)

Waiblinger S, Boivin X, Pedersen V, Tosi MV, Janczak AM, Visser EK and Jones RB (2006). Assessing the human–animal relationship in farmed species: A critical review. *Applied Animal Behaviour Science* **101**, 185–242.

Walker J, Dale A, Waran N, Clarke N, Farnworth M and Wemelsfelder F (2010). The assessment of emotional expression in dogs using a Free Choice Profile methodology *Animal Welfare* **19**, 75-84

Wang Z-Q, Porreca F, Cuzzocrea S, Galen K, Lightfoot R, Masini E, Muscoli C, Mollace V, Ndengele M, Ischiropoulos H and Salvemini D (2004). A newly identified role for superoxide in inflammatory pain Journal of Pharmacology And Experimental Thera

Webster J (2005). The assessment and implementation of animal welfare: theory into practice. *Scientific and Technical Review of the OIE.* **24,** 723-734.

Wemelsfelder F, Hunter EA, Mendl MT and Lawrence AB (2001). Assessing the 'whole animal': a Free-Choice-Profiling approach. *Animal Behaviour* **62,** 209-220.

Wemelsfelder F, Nevison I, and Lawrence AB. (2009). The effect of perceived environmental background on qualitative assessments of pig behavior. *Anim. Behav*. **78**, 477-484.

Wemelsfelder F, Hunter, AE Paul, E S and Lawrence AB (2012). Assessing pig body language: agreement and consistency between pig farmers, veterinarians and animal activists Submitted

Whay HR, Main DCJ Green LE and Webster AJF (2003). Assessment of the welfare of dairy cattle using animal-based measurements: direct observations and investigation of farm records. *Veterinary Record* **153**, 197-202.

Whay, HR (2008). On-farm Animal Welfare Assessments and Audits. In: Proceedings of the UK American Association of Bovine Practitioners Conference, pp.73-77.

Wichman A, Keeling LJ and Forkman B (2012). Cognitive bias and anticipatory behaviour of laying hens housed in basic and enriched pens. *Applied Animal Behaviour Science* **140**, 62–69.

Widowski T M and Hemsworth PH (2008). Housing hens to suit their 'needs'. In: Proceedings of the XXIII World Poultry Congress. World's Poultry Science Journal and the World's Poultry Science Association, The Netherlands, pp. 241.

Wolfensohn S and Honess P (2007). Laboratory animal, pet animal, farm animal, wild animal: which gets the best deal? *Animal Welfare* **16**, 117-123

Yao J, Burton JL, Saama P, Sipkovsky S and Coussens PM (2003). Generation of EST and cDNA microarray resources for the study of bovine immunology. *Acta Veterinaria Scandinavica Supplement* **9,** 89-103.

Yeates JW and Main DCJ (2008). Assessment of positive welfare: A review. *Veterinary Journal* **175**, 293-300.

Identify and Integrate Measures of Animal Welfare that Meet the Needs of Animals and Society – Part 2

Final Report
APL Project 2011/1036.421

February 2013

CSIRO Animal, Food and Health Sciences
Drewe Ferguson & Dr Ian Colditz
671 Sneydes Road
Werribee Vic 3030
Murdoch University
Teresa Collins
90 South Street
Murdoch WA 6150
Lindsay Matthews & Associates Research International
Dr Lindsay Matthews
The University of Melbourne
Animal Welfare Science Centre
Prof Paul Hemsworth
Faculty of Land & Food Resources
Vic 3010

Executive Summary

Part 1 of the Review described the three principle conceptual frameworks for assessing animal welfare: biological functioning, mental functioning and naturalness.

The design of a welfare assessment scheme is influenced by the purpose of the scheme, the method of the scheme's implementation including whether it is to be compulsory or voluntary, the processes for ongoing external verification of the scheme, how the scheme is to be funded and managed.

Challenges in designing a scheme include validation of welfare measures, weighting and scaling of different measures if a single score is to be generated, and whether trade-offs between different measures should be accommodated. For efficient implementation, the scheme should be parsimonious in the number and complexity of measures yet these measures need to be sufficient to address the scope of welfare concerns under assessment.

The three principal purposes for welfare assessment are: regulatory compliance, market assurance and welfare management. It may be difficult or impossible for a single assessment scheme to serve all three purposes.

From consideration of this background and the welfare assessment schemes in use around the world, we propose a Unified Field Index (UFI) for assessing welfare in commercial livestock enterprises. The UFI incorporates measurement domains that address key areas across the conceptual frameworks of biological functioning, mental functioning and naturalness. The UFI is a generic scaffold for application across livestock species and production systems. We have also considered and discussed how the UFI could be implemented within the Australian livestock industries. This was perhaps outside the original project brief but the proposed implementation process is important to consider because the overall utility of the UFI will be ultimately underpinned by the manner in which it is applied in practice. Moreover, there are some novel aspects associated with the implementation process such as welfare performance benchmarking.

The combination of the UFI and the proposed implementation process will create a tool for livestock owners and managers to monitor and manage welfare on their livestock enterprise. The implementation process will generate records for internal and external audits that provide evidence of the welfare performance achieved within the enterprise, and through external audit can ensure integrity of the welfare assessment process. The UFI together with the implementation process provide the basis for a new national livestock welfare performance program. It is recommended that the program also adopt risk management features that have proved successful for product quality assurance schemes used in other agricultural sectors in Australia.

The proposed Unified Field Index is comprised of 4 modules that address:

1. Animal based measures
2. Resource based measures
3. Management based measures
4. Other ethical criteria.

It is proposed that the UFI is implemented through a cyclical process operating at two levels:

Level 1 – Within the enterprise – conducted by the livestock manager

- Risk identification and assessment
- Intervention or corrective actions
- Monitoring of key variables
- Self audit (internal)
- Review

Level 2 – external processes – conducted by auditors and analysts

- External auditing
- Across enterprise benchmarking

A staged or incremental process of implementation may be preferred by livestock sectors where the initial effort is directed towards the development and implementation of Level 1 components on-farm. The Level 2 external processes could then be developed and introduced at a subsequent stage.

Insufficient data are currently available to set welfare performance benchmarks for Australia's commercial livestock industries. Therefore, across enterprise analysis of data acquired through the program is proposed as the basis for establishing benchmarks of good welfare management.

Oversight of the program could occur through an agency tasked with ensuring consistency in implementation of the program across livestock industries and consistency in data analysis and interpretation. The proposal aligns with the cross sectoral goals of the National Animal Welfare RD&E Strategy and the objectives of the Australian Animal Welfare Strategy (AAWS).

The UFI is a framework only. Specific details of measurements within each measurement class and subclass need to be developed. We suggest that this task be undertaken by each industry through consultation with relevant stakeholders in a process facilitated by a new oversight agency.

Strengths of the proposed UFI implemented through a national livestock welfare performance program include:

- Consistency of language for describing and interpreting welfare performance across livestock sectors.
- Integrity of welfare assessment provided through external audit and cross-sectoral oversight.
- Creation of welfare management tool for use on farm by livestock managers.
- The program is based on continuous improvement where refinements are incorporated through on-going scientific and industry validation of welfare standards.
- The program is suitable for generating knowledge about welfare standards through industry data then evolving into a product assurance or standards compliance scheme that are acceptable to government and society.
- Unlike previous welfare assessment schemes/indexes, extends the concept of good animal welfare to encompass a broader concept of good livestock management.
- Feedback through benchmarking enables establishment of attainable goals for ongoing improvements in welfare performance.
- Treats welfare as a continuous performance attribute like growth rate or milk production rather than a pass / fail judgement of an enterprise. In doing so, creates a culture for continual improvement of welfare performance.

- The assessment module addressing "Other Ethical Criteria" separates welfare performance as assessed though the first 3 modules from claims made about the ethical practices used in food production (e.g. organic, free of added growth hormones) that are currently conflated in the market place with animal welfare. Generation of this module is an ambitious goal that might be suitable for a later stage development of the UFI.

Weaknesses of the UFI include:

- Potential lack of confidence by consumers in a scheme based on self-assessment of welfare performance by industry.
- Reframing welfare in terms of performance rather than standards may face resistance from welfare advocacy groups. Clarification of the difference between the two concepts will be important.
- Might not adequately address some aspects of community concern about farm animal welfare.

1. Introduction

Part 1 of this review outlined how the assessment of animal welfare is influenced by:

- Information gained from measurements of the animals under assessment, their environmental resources and their management
- The conceptual framework brought to the assessment task, and
- The purpose for which the assessment is being undertaken

Currently, the principle conceptual frameworks brought to the task of welfare assessment are:

- Biological functioning – deviation from normality as evidenced through measures of behaviour, physiology, health and productivity
- Affective states – as evidenced through measures of abnormal behaviours, mental states (positive and negative feelings) and cognitive function, and
- Naturalness – as evidence by attributes of the animal, in particular normal behavioural repertoires, and by attributes of its environment

We consider that it is desirable for an assessment system to combine elements of all these frameworks in order to provide a broad-based assessment. Here we address ways of developing an assessment system for use in the field in commercial livestock production enterprises. We propose a Unified Field Index for application across all livestock species and production systems that is combined with an implementation process to provide 1) a risk management approach to assist livestock owners and managers achieve good welfare outcomes for their animals, and 2) an audit function for external verification of welfare performance of the livestock enterprise and to establish welfare performance benchmarks for the industry.

2. The Purposes of Welfare Assessment

The purpose for which welfare assessment is being undertaken has a major influence on the design of an assessment system and the way it is implemented. There are divergent views on the need for welfare assessment and the role that assessment serves. Even if we restrict the design of a welfare assessment scheme to addressing the way animals are farmed to yield products for human use, and set aside welfare assessment of animals used in activities such as research, companionship, education, rehabilitation of criminals and the unwell, and hunting, there still remains a diversity of purposes for which assessment might need or want to be undertaken.

Three principal purposes for assessing welfare of farm animals are:

1. To determine compliance with policy, law and regulatory standards.
2. To assure both consumers and non-consuming members of society that aspects of the welfare of animals not articulated within 1 are being met, for instance freedom to roam in an outdoor environment or absence of suffering.
3. To assist owners and managers to monitor and manage the welfare of livestock in their care and responsibility.

These purposes can be summarised as compliance, market assurance and welfare management.
A number of design and implementation issues flow from the purpose for which welfare assessment is undertaken. These include:

1. Should assessment be compulsory or voluntary?
2. Should assessment be made against externally validated standards or against benchmarks established by within-industry comparisons of performance?

3. Should the assessment program commence in a fairly short time frame or after a period of further research to better define and validate standards?
4. Should assessment be conducted by external auditors making site visits, by internal assessment, or a combination of both?
5. Who pays?
6. What sort of structure is needed to oversee the assessment scheme?
7. What are the functions of the oversight structure?
8. Should (and could) the scheme attempt to fulfil all purposes or be more restricted in scope?
9. Should the scheme endeavour to draw current assessment programs serving the same purpose under the one umbrella or should the new scheme be just another player in a fragmented market place?
10. How far along the supply chain from land and resource inputs on farm to product management in store should the scheme extend?
11. How parsimonious should the measures be: is it sufficient to account for say 80 percent of poor welfare with just a few measures or should we keep adding additional measures to chase ever diminishing additional accounting of welfare states?

Not all these questions can be explored within the scope of this review as they will be determined in part by political decisions; however some of the prominent considerations are explored below

2.1 Assessment of Compliance

A scheme to assess compliance of farm practices with regulations and standards of farm animal management and care is likely to require inspections by external auditors who assess on-farm welfare against externally validated standards. Standards may need to be sufficiently robust to withstand challenge in a court of law. Oversight of the scheme may require a statutory body or a body with responsibilities assigned to it by parliament. Funding might be by compulsory industry levees or by government budget allocation. On the other hand, some voluntary welfare assurance schemes with wide adoption by farmers such as Red Tractor (http://www.redtractor.org.uk/) in the UK are criticised by welfare advocacy groups as only providing assurance of compliance with standards and regulations, and that they permit unacceptable farming practices (Eg: http://www.ciwf.org.uk/what_we_do/labelling/standards_analysis_report.aspx).

2.2 Market Assurance

Many factors have led to the plethora of assessment programs for market assurance of welfare standards including the divergent views of what constitutes good welfare and the commercial opportunities for farmer groups and retailers to service niche markets catering to ethical aspirations of some consumers. Undoubtedly some assurance programs have been developed with a political motive to leverage change in farming practices through the market impact of retailers and brands. Indeed the complex mix of contested views on what constitutes appropriate or meaningful measures of welfare reflected in the diversity of assurance programs currently in place is one reason this review was commissioned.

The EU Welfare Quality program entered this type of socio-political environment in 2004 with funding by the European Commission. It was an integrated program involving 44 institutes and universities (representing thirteen European countries and four Latin American countries) with specialist expertise in animal welfare health and production, and took over five years to complete. The project aimed to accommodate societal concerns and market demands, to develop reliable on-farm monitoring systems, product information systems, and practical species-specific strategies to improve animal welfare. Throughout the project, efforts were focused on three main species and

their products: cattle (beef and dairy), pigs, and poultry (broiler chickens and laying hens). Welfare Quality does not appear to have displaced or stopped the development of other market assurance schemes in the EU such as Red Tractor, Real Welfare, and AssureWel. In part, the continuing emergence of new welfare assessment schemes illustrates the evolving nature of scientific understanding and consumer concepts of good welfare. A capacity for a welfare assessment or assurance scheme to evolve with new knowledge and changing societal attitudes is likely to be an important attribute of the scheme.

It is likely therefore that no new welfare assessment scheme for market assurance can have the authority or complexity to be accepted by all consumers. Nonetheless, there is the opportunity to provide clarity to claims of current welfare assurance schemes through a new assessment framework based on scientifically established standards that other schemes can be judged against. The method of implementation of the framework will influence its perceived integrity. If standards implemented though the scheme are higher than those mandated in Regulations and Standards and Codes of Practice it is unlikely that participation in such a scheme could be compulsory. The broader the reach of the assessment framework across species and production systems and across stakeholder groups from farmers to welfare advocacy groups to retailers, the stronger the perceived authority of the scheme is likely to be. Funding and governance of such a scheme are likely to be influenced by these second order objectives of the scheme.

2.3 Welfare Management

A third purpose for welfare assessment is to provide farmers with information that enables them to improve the management of the welfare of animals in their care and responsibility. Benchmarking provides a method of performance management that is used by many industries where data from individual enterprises is collected and analysed across enterprises to identify the range of performance achieved and the management strategies underpinning good performance. Benchmarking has been widely adopted in agricultural industries and its power as a tool for technology transfer, for engaging farmer participation and for raising production performance in Australian agriculture has been clearly demonstrated by the Cropcheck program (Lacy, 2011). Cross enterprise benchmarking has recently been applied to on farm welfare management in dairy cows (von Keyserlingk et al., 2012) and pigs (Keeling et al., 2012b; Leeb, 2011), and to welfare assessment of road transport practices (Gonzalez et al., 2012).

Benchmarking also underpins a number of market assurance schemes in the horticultural industries in Australia such as Freshcare (http://www.freshcare.com.au/). Such schemes are typically managed and paid for by industry and involve external audit processes as well as self audit of on-farm practices and of compliance with best practice guidelines. The schemes can provide risk assessment and risk management tools for producers, and technical support to help producers reach performance targets. A significant limitation is the perception of a conflict of interest when industries engaged in self-regulation and self-assessment, although this perception does not appear to encumber some self-regulating farming practices such as organic farming which is accredited through the organic certification entities (http://www.australianorganic.com.au/ and http://www.organicgrowers.org.au/) of the organic farming peak body Biological Farmers Australia (http://www.bfa.com.au/).

3. Ways of Constructing an Index

We can break the problem of constructing a system for assessing welfare down into 3 elements. These are:

1. concepts of what constitutes good welfare;
2. models of the biology of the animal and how interactions of the animal and its group with their environment affect animal biology, and
3. the process for implementing and evaluating the index.

The first two elements most strongly influence the choice of indicators for inclusion in the assessment index and the third element influences operationalisation of the assessment system and the extent to which the overarching goals of the welfare assessment have been achieved. Together, elements 1 and 2 can be considered to provide the model of animal welfare used in the assessment system. Part 1 of the review described the 3 conceptual frameworks of welfare in currency today.

There are generic challenges to the choice of indicators used in an assessment system that are not unique to welfare assessment. For instance, in the design of methods for assessing agricultural sustainability, Binder et al., (2010)have noted that assessment models should aim for parsimony and sufficiency, and account for interactions between indicators used in the model. Thus simplicity in the type and number of measures included in an index (parsimony) needs to be balanced against adequacy of the chosen measures to capture critical information needed to make a valid assessment (sufficiency) while interactions between measures that influence the interpretive outcome of assessment should also be included in the assessment protocol. These design requirements for an effective assessment tool apply equally to welfare assessment.

Before an assessment of the welfare of an animal or group of animals can be formulated, information and measurements are required on the animals, on the resources in their environment and on how they are managed. Some of the variables measured are quantitative while others are qualitative measures (e.g. categorical scores). Measurements made on the animal like body temperature, blood cortisol concentration, and growth rate are examples of quantitative measures on a continuous scale, while measurements like access to companion animals, presence of skin lesions or causes of mortality are examples of qualitative or categorical measures on a nominal scale. An example of another qualitative variable is the level and type of training undertaken by stockpersons.

3.1 Mathematical Modelling of Animal Welfare

A challenge in constructing an index is to decide how to combine the diversity of measures and choose which are most informative. A rapidly growing and highly specialised area of research that could contribute to this task is mathematical modelling of biological processes. This approach attempts to use mathematical representation of biological processes in order to predict changes in a system over time as input variables change. Such models can be used as research tools, for instance to identify where relationships between components of the system require better understanding, and when the model is well developed, as management tools, for instance in design and implementation of a vaccination program. If developed for assessment of animal welfare, mathematical models might need to be constructed for each species, and perhaps for specific production systems to accommodate the key relationships within each scenario to achieve the desired predictive accuracy. As a long term goal, it is desirable for mathematical models of welfare to be developed, and some groups such as the Animal Welfare and Behaviour group at the University of Bristol are active in this area (http://www.bristol.ac.uk/vetscience/research/awb/). While the assumptions that underpin a welfare model would in the first instance be influenced by the

conceptual frameworks of welfare outlined in Part 1 of the review, a mathematical model of welfare has the potential to enable the validity of assumptions to be tested in silico during refinement of the model. However, external validation remains critical to the success of mathematical modelling of biological processes. Amongst the many types of biological models under development, ecosystem modelling may provide the example closest to animal welfare in terms of the diversity of biological functions and human influences that need to be accommodated by the model.

3.2 Categorical Scoring Systems

A common approach to constructing welfare indices has been to identify the domains that are to be scored or measured and then assess each domain in turn. Scores can be summed across domains to provide an aggregate value as an indicator of welfare (a process that is sometimes termed as integration), or alternatively threshold scores that must be attained within each domain or within a specified subset of domains can be set. This approach typically uses scores ascribed to the animals and their environments by the assessor rather than being based on assaying physiological variables and assessing them against the normal range of the variable for the species or class of animal. Although objective and quantitative, the index tends to be based on visual scoring by trained assessors rather than measurement of biological variables. The Austrian Animal Needs Index was the first example of this type of index (Bartussek, 1999). Assessment criteria can be weighted so that some criteria have more impact on the overall assessment than others. Assessment domains and their weightings are influenced by the conceptual frameworks of welfare described in Part 1 of the review. Often the rationale for the weightings will also be based on ethical judgements.

3.3. Hazard Analysis and Critical Control Point (HACCP) Systems

The concept that some domains of an index must be satisfied for the overall assessment to be favourable has been used in other quality control process such as HACCP. This system involves analysis of hazards within a production process and measurement of parameters at stages of the process that are determined to be critical to the outcome. HACCP has been applied to animal health monitoring (von Borrell, 2000) and welfare assessment in abattoirs (Grandin, 2000). HACCP represents a simple deterministic model of a process. The monitoring of variables at the critical control points provides data for ongoing refinement of the HACCP model for the specific production process. At the outset, a HACCP model of animal welfare would need to be pre-determined by a conceptual model of welfare such as the conceptual frameworks outlined in Part 1.

3.4. Risk Assessment

Risk assessment is conceptually closely related to HACCP in that a causal linear relationship is assumed to exist between components of a process. Hazards are identified and the risk of failure of the component or process is assessed in terms of what can go wrong, how severe the consequences would be and how likely failure is to occur. The potential for risk assessment methodology to be applied to welfare assessment has been reviewed by the European Food Safety Authority (EFSA Panel on Animal Health and Welfare, 2012). Biological knowledge of the animal and its production system underpins the identification and assessment of the welfare risks the animal may experience. Risk assessment therefore represents one method for implementation of the conceptual frameworks of welfare described in Part 1 of the review. As we see in the HACCP process above, risk assessment needs to be integrated into a management system for welfare. Thus actions are required to remove identified hazards or to reduce risks of exposure to hazards, and to monitor the cumulative impact on animals of exposure to hazards in a manner analogous to critical control point monitoring. The Fresh Care program in the Australian horticulture industry and approved supplier certification in the Australian macadamia industry provide two examples from agricultural industries where risk assessment and risk management through corrective actions and monitoring of the

production system are combined with record-keeping practices, self auditing and external auditing to provide quality assurance of horticultural products and production processes.

4. Challenges in Constructing an Index

There are several significant challenges associated with the construction of a multivariate index to assess animal welfare. These include:

- Identifying which variables to include
- Defining the measurement weightings
- Identifying critical thresholds in the index or with specific measures

The selection of measures to include in an index will ultimately be based on the trade-off between the validity and repeatability of the measure and the practical considerations when recording it within the production environment. Statistically, validity has two primary dimensions; internal and external validity. Internal validity relates to the strength of the relationship between the measure and the welfare outcome or risk based on experimental evidence. External validity is concerned with how well this relationship holds when tested in the general population or under industry or "real world" conditions. Unfortunately, demonstration of external validity has not received the attention it warrants. However, there are notable examples such as the work of Dawkins et al., (2004) and Jones et al., (2005) where the external validity of stocking density as a welfare input measure in broilers was examined.

From Part I of this review, it is recognised that in the context of welfare assessment, demonstrating the validity of welfare measures is an ongoing challenge. However, this should not be used as an argument for slowing the progress towards the development and implementation of on-farm welfare assessment systems.

The balance between input or resource and output or animal-based measures in an index will be governed by the production system, species and practical considerations. As discussed in Part I, whilst animal-based measures generally provide more accurate indications of the animal's state and have more universal utility (i.e. can be independent of the production system), they are often more costly and difficult to collect on-farm, particularly in extensive livestock systems. Consequently, resource or management-based measures are often favoured due to their practical advantages (Main & Webster, 2011). Defining the appropriate weightings of the index measures is another significant issue. For example, at the herd or flock level, what weighting should be given to measures that reflect expression of natural behaviours compared with incidence of disease or ill-health? The weighting for each may also vary depending on the production system (eg. the contrast between caged versus free-range egg production). In the case of selection indexes used in animal breeding, the weighting assigned to the component traits within the index are typically predicated on the economic value of each trait. For traits such as growth rate or milk yield, establishing the economic value is relatively straight forward using actual market data. Unfortunately, it is far more challenging in the context of animal welfare. Returning to the example above, whilst it will be possible to derive economic values for disease incidence it is more difficult for the expression of natural behaviours. Indeed, this may be even more problematic in the case of resource-based or input measures. However, there are examples such as the Austrian Animal Needs Index where weightings were derived for specific resources (Bartussek, 1999).

It is inevitable that the weightings for some index measures will need to be estimated and in order to make informed judgements, the application of risk assessment principles may be of use. For

example, if we understand the primary welfare risks within the production system (*hazards*) and can estimate their impacts on the animal (*severity* – 1 (low) to 5 (high)) and the probability of their occurrence (*probability* – 1 (low) to 5 (high)) it is possible to numerically rank each hazard based on the calculation of *severity x probability*. Higher weightings would then be given to those measures that reflect hazards or welfare risks with a higher overall risk rating and vice versa. This might be combined with economic values where these are known.

Whether it is relevant to a specific measure or a multivariate index, one of the most challenging issues is the definition of critical thresholds. Specifically, thresholds that delineate when management intervention is required to mitigate a potential poor welfare outcome occurring or the worst case scenario, when welfare has been compromised. The other complicating dimension here is the level this is applied – at the animal or herd or flock. Take for example the assessment of dairy cow lameness, what is the relative significance in terms of welfare when different proportions of the herd (1%, 5% or 10%) manifest clinical signs of lameness?

For some animal-based measures (e.g. body temperature), the normal reference values/ranges from clinical veterinary data can be applied to identify critical thresholds. Another approach is based on consensus of expert opinion also known as the Delphi method (Whay et al., 2003). However, for most cases, the setting of specific thresholds will be based on a blend of experiential, intuitive and pragmatic inputs. There are clearly other precedents outside the field of animal welfare where this has occurred. The setting of maximum vehicle speed limits for example, is a good case in point. When these were legislated early last century, there would have been lack of empirical evidence to support them. Now we have very good data to show the effects of vehicle speed on both collision incidence and personal injury (e.g. www.officeofroadsafety.wa.gov.au/campaigns/speed_august2006/index). As a consequence of this new evidence, the maximum limit in built-up or suburban areas was reduced from 60 to 50 km in the majority of Australian states and territories late last century. Similarly the principles of continuous improvement should also be applied in the setting of critical thresholds and in the evolution of welfare indexes.

5. Examples of Field Indices

The main examples of field indexes were described in Part 1 and key aspects of the major schemes are revisited below. To summarise, traditionally such indexes have utilised input-based measures to assess or assure welfare. As discussed, input measures are relatively easy to observe and record and, therefore, to assess the level of welfare compliance. However, unless environmental/physical parameters accurately predict welfare status (which most do not) they suffer from the serious disadvantage of not measuring the welfare of the animal directly. For example, animals under equivalent housing conditions may be in vastly different welfare states depending on the quality of the management or variation in ambient conditions. Thus, in contrast with input-based indexes (such as the Austrian Animal Needs Index or the RSPCA's Freedom Foods assessment system), the focus for the future should be on those systems that concentrate on using outcome or animal-based measures. The EU Welfare Quality protocols are the best-known and most comprehensive of the schemes that use (primarily) animal-based measures including. measures of health (e.g. injuries, diseases for health status), behaviour (e.g. panting for thermal comfort, qualitative behavioural assessment) or physiological state (e.g. body condition for hunger levels). Welfare Quality protocols use some resource-based measures (e.g. numbers of waterers for thirst levels, or reported use of anaesthetics for pain control). The Bristol Welfare Assurance Programme (http://www.etschool.bris.ac.uk/animalwelfare) also uses, in the main, animal-based measures. Apart

from some question marks about the validity of some measures, from a practical point of view, Welfare Quality has been criticised for not being readily implementable on farm.

The current focus in Europe with the development of practical assessment systems has now shifted to identifying as few as possible animal-based measures that will adequately describe the welfare status of the animals. In this way, there is a higher likelihood that the measures can be implemented cost-effectively by producers, with the result that the animals themselves will reap the benefits. Arguably, the furthest-advanced system of this nature is AssureWel (http://www.assurewel.org/) a joint project between the UK's RSPCA, University of Bristol and the Soil Association and which is being introduced to the RSPCA Freedom Food and Soil Association assurance schemes for all the main livestock species. AssureWel is a 5 year (2010-2015) project that has identified a list of species-specific measures and has assessment protocols in use for laying hens and dairy (see Appendix 1 - dairy cow assessment protocol). Assessment protocols for pigs, broilers, beef cattle and sheep are currently being field tested. Producers are given feedback and support (including benchmarking) to help monitor and improve welfare on their farm.

The British pork industry is evaluating and introducing a parallel system called Real Welfare (http://www.bpex.org.uk/R-and- D/welfare/realwelfare.aspx) for pork production.

The Real Welfare scheme utilises a very limited set of indicators. For finishing pigs they are: tail lesions, body wounds, lameness, enrichment use, in need of hospital pen. For sows it is: shoulder lesions, vulva lesions, body condition, body wounds, lameness, enrichment use, in need of hospital pen. The practicality of these measures has been evaluated in on-farm studies. The pork industry appears to be using the results for benchmarking and development of herd health plans with veterinarians, but there is an appreciation of the need to identify acceptable standards as well. The thresholds for (un)acceptable welfare are being defined by reference to expert opinion.

The use of selected animal-based indicators has proved useful in improving welfare of livestock at slaughter (Grandin, 2010, 2012). It may well be advisable to include some resource-based measures in field indexes, as these provide useful indications of risk and, thus, ways to identify and ameliorate unacceptable hazards (EFSA, 2012).

5.1. Health and Production Modular Indices

Other scoring systems using a specific subunit or modular index have been developed for specific use and in particular contexts of interest. The merits of several methods of assessing welfare at the herd level have been reviewed by Johnsen et al., (2001) and include using farm records or animal observation. The use of outcome based measures of poor animal health, using farm records could potentially identify farms that use management practices that place animals at risk. For example, Vasseur et al., (2012) compared mortality and morbidity levels and routine calf management practices on 115 dairy farms across Canada, Austria and Germany. Although some farms showed higher juvenile mortality incidence than others and used more management practices that were recognised as putting calves at risk, conclusions were problematic because 1) measures only give an indication of health during a narrow time window and 2) measures of incidence require the use of accurate farm records. Record keeping may be far from ideal and producers themselves may have difficulty estimating calf mortality or morbidity on their farms and the mortality and morbidity records available may not reveal the extent to which at risk management practices are used.

The use of direct observations of dairy cattle in combination with the investigation of farm record data has been trialled to assess welfare (Whay et al., 2003). Through a process of consultation with

experts, a detailed assessment protocol based on observable signs such as lameness, coat condition, hock injuries and body condition score was developed and tested on 53 UK dairy farms (median herd size was 108 cows). The farms were ranked by 50 experts from 1 (best) to 53 (worst) for each measurement, and the overall rank of each farm obtained by calculating the mean of all the measurements on that farm. There was broad agreement among the experts on the most serious problems (lameness and hock injuries) and that a percentage of farms needed to take action to reduce the incidence of mastitis and lameness. Such agreement supports the value of observation based assessment of animals but further work is required to limit the number of observations needed and reduce time taken per visit.

5.2. The Assessment of Body Language: Qualitative Behavioural Assessment (QBA)

The major limitation of the Welfare Quality protocol is the time consuming nature of farm visits and the lack of transparency in how scores are integrated into welfare outcomes. The evaluation of an average sized dairy farm (200 cows) takes around seven to eight hours by trained staff. Although part of the Welfare Quality protocol, qualitative behavioural assessment (QBA) could be also be used as a screening tool or index to identify farms with compromised welfare. As stated in part 1 of this review, QBA relies on the observers' ability to integrate details of the animals' demeanour and its context, using a whole animal approach. Observers use descriptors such as 'relaxed' or 'anxious' that reflects the emotional state of the animal. The advantages of using QBA as part of a welfare index is that it is not very time consuming – on farm assessment takes at most 30 min and it strongly focuses on the animal, thus it obviates the need to merge scores on different aspects of welfare. QBA has been shown to correlate well with the animals' physical state, and has high inter-observer and intro-observer reliability despite observers having different cultural backgrounds and different levels of experience in animal behaviour (Napolitano et al., 2012). There is also emerging evidence to demonstrate that it's validity as an indicator of affective state in pigs (Rutherford et al., 2012). This is significant given the growing concern from the community about the treatment of animals centres around how the animal is feeling. Hence, including QBA in an assessment index is likely to receive strong support from the community.

5.3. Welfare Auditing Programmes at Slaughter

Since 1999, auditing programmes that utilise five numerically scored criteria have been used successfully by major restaurant chains to monitor animal welfare at US beef and pork slaughter plants. In order to remain on the approved supplier list, plants need to meet certain percentage scores (Grandin 2012). The audit has five numerically scored criteria that are called critical control points, which cover both resource and animal-based measures, and are fully described by Grandin (2010):

1) percentage of animals that slip or fall down during handling
2) percentage of animals moved with an electric prod
3) percentage of animals vocalising in the stunning box or restrainer
4) percentage of animals stunned effectively with one application of the stunner and
5) percentage of animals rendered insensible when hoisted onto the bleed rail (has to be 100% to pass the audit).

A minimum or maximum percentage is required on all five of the numerically scored criteria. For example, for vocalisations, less than 3 and 1% of cattle and pigs vocalising, respectively is required.
This auditing system uses the same approach as the Hazard Analysis Critical Control Point (HACCP) approach in food safety. The principle is to use a few outcome measures that can detect a variety of problems. Although it only uses a few points to measure welfare it satisfies the practical constraints

of time and costs. Today, all plants that supply US McDonalds and Wendy's has this type of audit by a third party company every three years.

Baseline data collected in 1996 before the restaurant audits started indicated that only 30% of beef plants could stun 95% of the cattle with a single captive bolt shot. In 2010, all 32 audited beef plants achieved this standard (Grandin 2012). Most plants were able to pass the audits without having to invest in capital equipment, but intensive programs of employee training were required. The audits have resulted in great improvements because plants that do not improve are removed from the approved supplier list.

Other parameters like presence of skin lesions or meat quality (percentage of dry, firm, dark meat) have been identified as being a 'fast' and relatively easy method to assess animal welfare at the abattoir. Combining such post mortem measures with other on–farm measures (productivity) may represent a potential integrated welfare tool. Meat with a higher percentage of DFD has been associated with fatigued and stressed animals, and those having experienced longer transport journey (Fabrega et al., 2007).

5.4. Using Breeding Objectives to Improve Welfare

Traditional breeding practices rely on the definition of a breeding objective which identifies traits that are economically important based on market requirements. Selection of breeding animals that meet these objectives are made by measurement and genetic evaluation. Historically, breeding objectives have focused on productivity (growth, milk yield, feed efficiency) and functional traits (health measures, fertility). The weights applied to traits in the objective reflect their economic value to the producer and these in turn must be known or estimated (Lidauer et al., 2000). However, as societal concerns about animal welfare have centred on the impact of intensive environments and management practices on the animal, practical contributions towards reducing welfare problems can be made by creating more balanced breeding programs (Lawrence et al., 2004). Societally important traits may have an economic and a noneconomic value and are sufficiently heritable for effective genetic selection, although many genetic and phenotypic (co)-variances still have to be estimated. Yet often is not clear to a pig or cattle breeding organization how it may deal with the non-economic value of breeding-goal traits. In addition, new approaches are needed that estimate the value of non-market values issues such as the pain or discomfort associated with lameness.

Kanis et al., 2005 described a retrospective selection-index method to obtain the proper weights for societally important traits in the breeding goal for pigs. The method offers the possibility to estimate the societal costs of selecting for economic traits only, as well as the societal benefits and the economic costs of selecting for traits with a non-economic value. Traits such as temperament, stress resistance and robustness have sufficient genetic variation for successful selection and considered good candidate goals. This shows promise as a useful tool for the development of novel sustainable breeding goals.

Breeding goal definition is a research area of ongoing interest; as knowledge on modelling (both single and multiple trait regression) is improving, and production circumstances are continuously changing (Lidauer et al., 2000). Finding the proper balance between genetic progress of all breeding traits requires some insight into the future importance of the various traits. Such predictions can be based on market and societal trends with respect to consumer preferences and societal concerns (Kanis et al 2005). Thus, the applications of approaches that require ethical priorities in the aggregate genotype are likely to contribute to sustainable production systems.

6. Requirements for a Single Index

An ideal scheme might be to establish a set of minimum requirements that can satisfy many objectives; a) assures the well-being of animals b) meets market and regulatory demands in terms of certification and c) allows different standards of welfare products to be compared. However for such a program to be feasible, it must meet practical constraints in terms of time and resources and must not require complex training for those performing the audit. Therefore, it may be difficult to develop a "one size fits all" scheme. The core challenge is to identify a range of parameters that are easy to measure on-farm, inexpensive and non –invasive, that are scientifically credible and best represent animal well-being, that have meaningful thresholds when applied but that also satisfy consumer or community concerns. As no gold standard for objectively determining welfare exists, the need for collection and integration of data using different parameters is essential. In addition, to reach consensus across industries on any single welfare index, a number of issues must be addressed.

General challenges:

1. Parameter/measurement selection that satisfy various stakeholder groups and achieve specific objective
2. Parameter/measurement weightings and aggregates (within domains, across domains, across lifetime of animal)
3. Qualifications and experience of experts to design and apply models
4. Defining thresholds for different levels of welfare for certification and QA programs
5. Measurement at individual or group level

Specific challenges:

1. Lack of measures for some states
2. Lack of validation, sensitivity and specificity of many measures
3. Lack of practical and easy to use measures.
4. Encouraging uptake with producers and dialogue with large retailers to ensure that retailers value welfare initiatives in order to maximise any competitive advantage brought to producers
5. Balancing parsimony in the number and complexity of measures with their sufficiency to assess welfare

Welfare assessment by definition involves integration of different parameters; integration can be done different ways and inevitably involves human judgement. Identifying the structure behind the integration process will highlight the advantage and disadvantage of each approach and also make transparent the points at which human judgement or ethical views are required. Methods to assess welfare vary in their applicability to meet specific goals. Some methods will be better used in assisting the individual farmer to improve the welfare of animals within his production system and yet others will be able to compare productions systems well. Achieving good validity for the selected measures requires experts to be clear on the overall objectives of the assessment program.

7. Our Recommendations for an Index

Based on the above considerations, we recommend a Unified Field Index that incorporates risk management, auditing and benchmarking functions. The index provides a framework of assessment domains that are combined with a process that in combination delivers a tool for livestock managers

to monitor and improve the welfare of animals in their care, and generates records for internal and external audits that can provide assurance of the welfare performance achieved by the livestock enterprise, and ensure integrity of the assessment process.

The Unified Field Index provides a framework that should be applicable to all livestock enterprises. The index provides a suite of ***domains*** that are actualised through a ***process***. The Unified Field Index together with its process of implementation have the potential to provide a new national livestock welfare program.

The process is cyclical and follows a sequence operating at two levels:

Level 1 – Within the enterprise – conducted by the livestock manager

- Risk assessment
- Corrective actions
- Monitoring of key variables
- Self audit
- Review

Level 2 – external processes – conducted by auditors and analysts

- External auditing
- Across enterprise benchmarking
- Review of benchmarks, and over time translation of some into standards

Typically, the internal component of the cycle would be completed once per year, although some production systems might be better suited to longer or shorter cycles. Records of the internal process provide a basis for external audits to ensure integrity to the welfare management and assessment system, as well as data for across property benchmarking. Benchmarking provides the important functions of establishing performance levels for typical and leading enterprises within the industry, and performance goals and management strategies for low performers to emulate.

Implementation in the first year might be onerous for the livestock manager but in subsequent years should become less demanding as monitoring processes and recording systems become better established. A staged or incremental process of implementation may be preferred by livestock sectors where the initial effort is directed towards the development and implementation of Level 1 components on-farm. The Level 2 external processes could then be developed and introduced at a subsequent stage.

Some of the data required for welfare benchmarking is already collected by participants in genetic improvement programs managed through programs such as the National Beef Recording Scheme, Sheep Genetics and Australian Dairy Herd Improvement Scheme. New ways of describing and interpreting data on a whole of herd or whole of flock basis provide the potential for monitoring the welfare performance of enterprises. In addition, genetic analysis of the data collected through the program might be suitable for estimation of EBVs for new welfare traits.

Veterinary medicine relies on reference ranges for clinical measures like cortisol, heart rate, rectal temperature, blood metabolites, hormones etc when these variables are measured to aid in diagnosis of disease. Reference ranges for most production and welfare variables are not currently available for Australian livestock enterprises. An important function of benchmarking is the establishment of reference ranges for welfare related variables which will provide the basis for assessment of welfare performance of Australia's livestock industries.

The assessment domains used in the Unified Field Index are grouped into three modules, plus one optional additional module.

The modules are:

1. Animal based measures
2. Resource based measures
3. Management based measures
4. Other ethical criteria.

Within each module the domains represent high level headings which describe areas for assessment that we consider to be generic to all (or almost all) production systems and species. For instance, the Animal Module has 5 domains: behaviour, health, affect, production, reproductive performance, and holistic attributes of the animal. Each Domain is subdivided into a classes and subclasses.

The components of the Index are presented in Appendix 1.

The domains attempt to represent the key areas of importance within each of the 3 conceptual frameworks for assessment of animal welfare outlined in Part 1 of the review. The Unified Field Index provides an advance on previous indexes and protocols for assessing welfare by inclusion of a range of management practices and operator skills within the assessment of the welfare performance of the enterprise as well as the more commonly included animal based and resource based measures. Details like suitability of the animal genotype for the production environment and impact of breeding objectives on welfare standards of future generations have also not been included in previous welfare assessment schemes. More details on assessment domains are provided below.

A uniform architecture of the Unified Field Index and a standardised implementation process are proposed so that a consistent concept of livestock welfare and a consistent language for describing livestock welfare performance can be used across all Australia's livestock industries.

Oversight of the program will be important to its success. Important functions for the body tasked with oversight of the program include:

1. consistency in data analysis and interpretation across livestock industries
2. data integrity
3. consistency in implementation in different livestock species
4. external reporting

We do not envisage that the oversight body be engaged in development of welfare policy or welfare standards or be a spokesbody on such issues. This proposal aligns with the cross sectoral intent of the National Animal Welfare RD&E Strategy and the objectives of the Australian Animal Welfare Strategy (AAWS).

The structural elements, information flow and applications of the UFI and associated implementation process are illustrated in Figure 1.

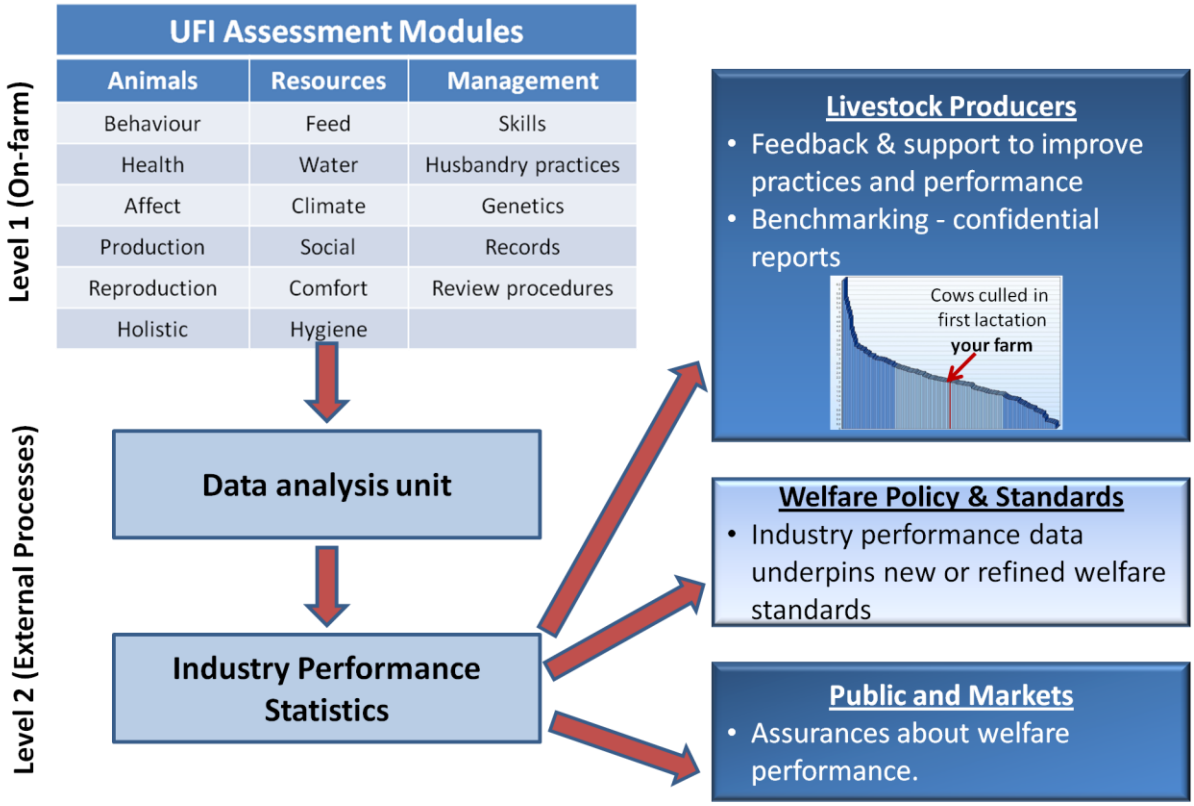


Figure 1: UFI structure, information flow and applications

7.1. Validation and Implementation of the Welfare Performance Program

The Unified Field Index provides a structure around which a welfare management and assessment tool for each industry or production system within each industry can be built. We see the development of the details of the measurements needed within each domain to be a task to be driven by each industry, in consultation with all stakeholders and animal welfare scientists. We see this as a means to engage industry in ownership of the welfare management and assessment tool for its industry. The process for implementation on farm is also designed to engage each livestock manager in ownership of management of animal welfare in their enterprise, through assessing risk, risk reduction, monitoring, record keeping and periodic review. If well tailored to each industry or production system, the new welfare program should also provide a valuable educational tool that through monitoring and assessment methods can help managers understand and improve animal welfare within their enterprise.

These tasks of risk assessment, risk reduction, monitoring, record keeping and periodic review are already performed on most well managed enterprises. A substantial challenge in developing the welfare management and assessment tool is to find a way to minimise the burden of compliance on livestock and business managers. A high priority goal should be to develop a system that can be integrated with other records needed for compliance with regulations such as chemical use, vendor declarations and occupational health and safety and ideally also with other livestock recording systems, product assurance programs and business records management systems.

What we present is a concept for a welfare management and assessment tool that could provide quality assurance of welfare **performance** for livestock industries. We don't expect that we have thought of or are familiar with all the issues around development of QA systems and how to best implement them. We therefore recommend a workshop with industry representatives and experts

in development of industry assurance programs and in the design of data collection and analysis systems for genetic improvement programs, together with providers of livestock management software tools, welfare advocacy groups and retailers to explore the feasibility of developing a system that can be integrated with other QA and livestock management systems.

A key issue for consideration is the process of implementation. A staged or incremental process of implementation may be preferred by livestock sectors where the initial focus is on the development of the UFI protocols and risk assessment modules on-farm. The subsequent stages of benchmarking and external auditing could be introduced at a later date.

The optional Module 4 is included with this goal of integration in mind as a means for incorporating other ethical claims of animal based production systems within a single framework. Further development of this module is needed. Some subclasses of this module are already the basis of ethical claims about animal based food products such as hormone use (e.g. free of added growth hormones), confinement practices (e.g. free range, barn laid) and production philosophies (e.g. organic, biodynamic). With the possible exception of confinement practices, these are ethical claims rather than welfare assurances, and compliance with the standards required by the certifying body for each of these ethical practices does not imply that welfare performance benchmarks assessed through Modules 1 to 3 are otherwise satisfied. Hence these ethical claims are separated from the welfare management and assessment criteria in modules 1 to 3.

The power of benchmarking as a tool for technology transfer, for engaging farmer participation and for raising production performance in Australian agriculture has been clearly demonstrated by the Cropcheck program (Lacy, 2011). The substantial experience within the Australian agricultural research and extension community in implementation of benchmarking programs should assist in the design and implementation of the welfare performance program. The utility of benchmarking prior to the setting of performance standards has also been highlighted through the experiences of the Welfare Quality Project. When the initial standards established by the WQ reference panel were tested in a range of dairy farms, very few farms passed indicating the bar had been set too high (Matthews *pers. comm*.).

7.2. Interpretation of Welfare Performance

While data is available from commercial livestock enterprises in Australia on many of the variables identified in the UFI, little has been analysed and interpreted in ways that provided information on welfare performance. Thus it is uncertain what values should initially be set as indicators of good welfare performance. To address this deficiency, cross enterprise analysis of data collected through the welfare performance program is proposed as a method to establish reference ranges for welfare performance currently achieved by commercial producers. From this analysis, benchmarks can be established for good welfare performance, and attainable goals set for improvement of performance by low ranking enterprises. As with other benchmarking programs in agricultural production, these performance goals will be supported by information on the management practices used by high performers and can be used by industry to lift overall welfare performance. Thus performance recorded through the welfare program provides feedback to producers and industry that can stimulate higher achievement in subsequent years. The power of this approach to improve welfare outcomes on farm has recently been recognized (von Keyserlingk et al., 2012). The cross sectoral role of the oversight body could provide a degree of autonomy and uniformity to the process of data analysis and interpretation that should strengthen the integrity of the program.

The structure of data acquired through the welfare program and the challenges in analysing this data are likely to be very similar to the data analysis tasks addressed in livestock genetics improvement programs. It is likely that expertise within genetics groups servicing Australia's livestock industries could make a highly valuable contribution to achieving the goals of the program.

7.3. What's New with the Unified Field Index?

The UFI represents a philosophical shift from past welfare assessment procedures on several counts. The UFI replaces the concept of welfare standards with the concept of welfare performance as a continuous trait (or suite of traits) or continuous attribute of an enterprise. Standards are accompanied by the risk that they are interpreted as all or nothing thresholds that only need to be exceed but not continually improved upon. Welfare performance is measured though benchmarking of performance across enterprises within an industry. Thus the program measures what is achieved and what can be achieved in real word circumstances. Though not yet substantiated by any evidence, it seems likely that welfare performance will be more easily linked to economic performance than standards are, as several of the components of welfare performance such as disease and growth targets are drivers of economic performance (von Keyserlingk et al., 2012).

The UFI introduces breeding objectives and genetic management practices into the assessment of welfare performance.

A significant feature of the concept is that performance is owned by industry rather than imposed on industry in the way standards tend to be. Ownership of welfare performance is likely to improve the incentive for improvement (Leeb, 2011).

The UFI has the potential to be a voluntary scheme starting from the current imperfect knowledge base to provide a management tool for improving welfare performance. Some of the benchmarks might have the potential to evolve into standards through external scientific and industry validation (Keeling et al., 2012a). The scheme might be suitable to management on a (semi) commercial basis similar to genetic improvement programs with seed funding from government and industry RDCs. As the scheme matures it might be able to provide market assurance and regulatory compliance functions through assessment of performance against the evolved standards.

These changes from contemporary welfare assessment schemes provide strengths and weakness. Important weaknesses include lack of external validation of welfare performance against welfare standards especially during early phases of the program, and the perceived conflict of interest for industry in assessing its own welfare performance. There could also be a fear within industry that data collected on welfare performance might be used against industry although this threat accompanies other types of welfare assessment programs (Croney and Anthony, 2010). Continual growth in the number of producers reporting use analgesia for mulesing on the national wool vendor declaration form might provide evidence against this fear http://images.wool.com/pub/flystrike_9_Grave_NWD_0610.pdf accessed 3/2/2013).

The UFI does not at this stage provide a mechanism for integrating independent measures into an aggregate or weighted score.

7.4. Assessment Domains Included in the Unified Field Index

Our vision is that the assessment domains would be largely consistent across species and production systems, although some minor modifications of the domains between species may be needed. Fine tuning of the Unified Field Index to the nature of individual livestock species and their production systems would largely occur at the level of the class and sub class of measures within each assessment domain. This tailoring to species and production system is analogous to the hierarchical design of the EU Welfare Quality program in which the 4 welfare principles (good feeding, good housing, good health and appropriate behaviour) are manifested through 12 welfare criteria which in turn are assessed through a suite of measures that are designed to be appropriate for each livestock species, as illustrated for dairy cattle in Figure 1.

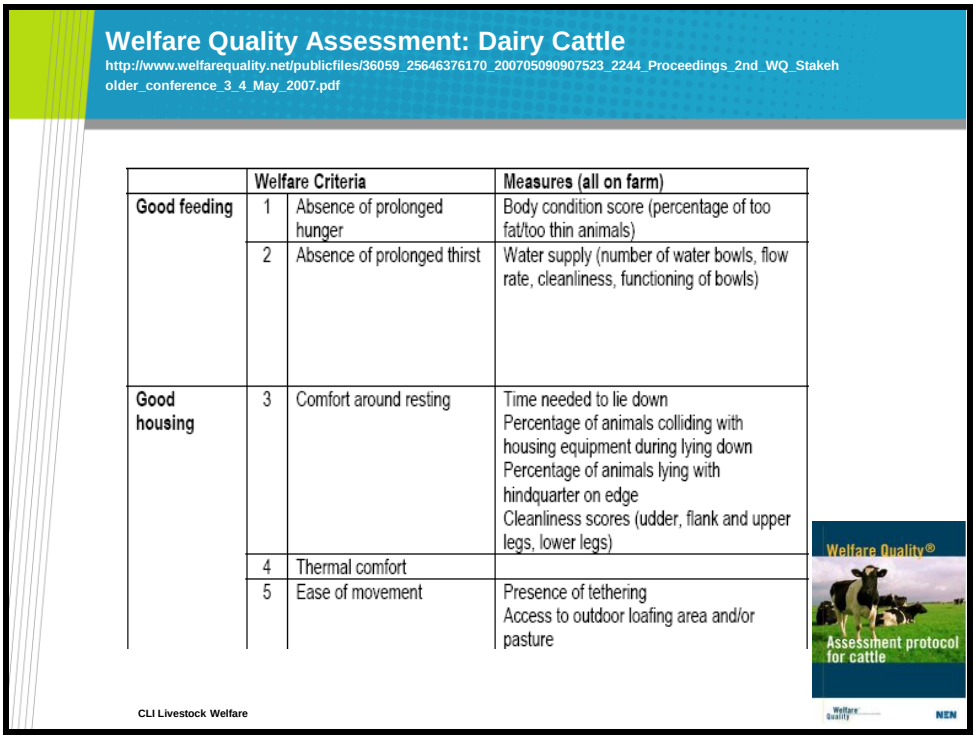


Figure 1: Welfare Quality Assessment protocol: Dairy Cattle

A number of the proposed classes and subclasses within the assessment domains are drawn from clinical veterinary medicine and animal production. Initial assessment of the health of an animal begins with visual inspection of the animal and its environment before hands on assessments are made. Many production variables are sensitive to stress and health status, and analysis of these across the whole flock or herd with particular emphasis on outliers and deviations from performance targets can provide information on welfare performance. Research studies on welfare usually combine these types of measures with more invasive and intensive measures of health, physiology, behaviour and production that are not readily adapted to field monitoring of welfare performance on commercial livestock enterprises. The dictum "you manage what you measure" underpins a number of the proposed new measurement domains within the Unified Field Index for assessment of welfare.

7.4.1. Animal Module

- Behaviours

Abnormal behaviours: Animals display a range of abnormal behaviours when stressed or when exposed to adverse environmental conditions such as heat. Stereotypic (repetitive, rhythmic) behaviours are common in intensively housed animals with impoverished environmental conditions. Fearfulness and other indicators of adverse reactivity of animals to the presence of humans that can be indicative of poor animal handling practices sits in this class.
Social behaviours: Social behaviours like mutual grooming and play are important indicators of the social health of social species like cattle, sheep, pigs and chickens.
Self care: Self care is reduced in animals stressed by poor nutrition, ill health and social bullying.

- Health

Mortality: Raw data on mortality can be criticised as being a worst case outcome of welfare management. Nonetheless, deaths do occur in well managed livestock enterprises and the information that death rates provide about welfare management should not be ignored by livestock managers. Data on the numbers of animals found dead versus animals euthanased is a valuable indicator of the level of monitoring and intervention to prevent suffering in moribund animals.
Morbidity: Disease compromises welfare and data on disease prevalence is central to effective disease control and good welfare management.
Current status: This domain provides the basis for visual assessment of the general health of the individual and the group. These measures are examples of ones that would be used in external audit for a snapshot of the flock or herd as well as during ongoing monitoring by the livestock manager.

- Affect

Validated measures of affect for application on farm have not been developed and further work is needed in this area. Nonetheless, the assessment of demeanour or behavioural expression as used in clinical assessment of animals by veterinarians and by pen riders in feedlots are two examples of visual assessments that probably provide information on the emotional status of animals associated with ill health that could be used as indicators of affect during the early stages of development of the new welfare program. Standardisation and refinement of these measures and development of new field based measures of affect will be needed.

- Production

Analysis of performance against production targets, especially the identification of the number of outlier animals, provides information on the level of management of the enterprise. Increased risk of disease and death in animals below the group mean for growth, body weight and body condition is recognized in a range of species including lambs (Hatcher et al., 2008), dairy cows (Koeck et al., 2012), pigs (Fahmy and Bernard, 1971) and rainbow trout (Janhunen et al., 2012). While by itself performance against targets is not always informative of welfare, in combination with other measures it can help provide a picture of welfare management on the enterprise. The value of these data is greatly enhanced through benchmarking.

- Reproductive Performance

Reproductive performance is highly sensitive to stressors and diseases that affect welfare, as well as being fundamental to the profitability of the breeding enterprise, and is strongly influenced by the quality of management that animals receive and the suitability of the genotype for its environment. This class could be included with Production, but is separated to emphasize the very high importance of good management of reproduction for achieving good welfare outcomes.

- Holistic Measures

This assessment domain is included to acknowledge that some production systems and production philosophies value whole of animal attributes of welfare. We are not sure how these might be measured. Ability of turkeys to reproduce by natural mating might be an example of a holistic measure. In the absence of appropriate measures, this domain could be deleted.

7.4.2. Resource Module

- Feed

Feed quality and quantity and suitability for the metabolic needs of the animals are important resources the animal needs.

- Water

Quality, quantity and ease of access.

- Climate

Range of climatic variables animals are exposed to.

- Social Resources

Social behaviours are listed in the animal module. This domain addresses social conditions that are not controlled by the animal such as stocking density, group structure and access to companion animals.

- Comfort

This domain addresses indoor and outdoor infrastructure including bedding, availability of shelter and protection from climatic extremes that affect comfort of animals.

- Hygiene

Hygiene conditions influence exposure to disease pathogens and environmental organisms that influence health and welfare.

7.4.3 Management Module

While access of animals to the environmental resources listed above is controlled by management practices, the Management Module addresses non resource aspects of animal management as well as business practices that impinge on animal welfare. It also addresses some of the process elements that are necessary for the program to be effective.

- Skills

Addresses whether stockpersons have appropriate skills training (or experience) and appropriate attitudes for working with animals.

- Husbandry Practices

This class addresses the methods used for husbandry practices, ages when the practice is performed, whether analgesia is used for painful procedures. Methods used for euthanasia are documented also.

- Genetics Management

The impact of genetic practices on welfare of livestock has not been addressed in previous welfare assessment schemes, yet many livestock welfare problems are closely linked to the genetics of the animals. So not only do genetic practices contribute to many welfare problems, they can also

provide a partial solution to many welfare problems. This heading addresses issues such as suitability of the genotype for the production environment, ways animals are selected for breeding purposes including use of EBVs for welfare traits within quantitative breeding objectives, and non-quantitative criteria (e.g. visual classing criteria) used for selecting breeders. Culling criteria, culling numbers and culling age provide important information on welfare performance of the enterprise, as defects and poor performance of individuals are often caused by or related to poor welfare.

- Records

Records are essential for self audit and internal review, as well as for benchmarking and external audit, and are also required for compliance with regulations around chemical use, vendor declarations and occupational health and safety.

- Review and Action Protocols

Success of the program depends on periodic internal review of data and processes. Review process, outcomes and actions taken need to be documented.

7.4.4. Module for Other Ethical Criteria

Modules 1 to 3 address welfare practices and could provide the basis for accreditation of the ethical standards of animal welfare in an enterprise. There are many claims made about the ethical practices used in food production that are conflated in the market place with animal welfare. This module separates these other ethical claims from ethical claims based on welfare practices addressed through modules 1 to 3. What is presented in module 4 is very much a first pass at development of such a module. The most important element of module 4 might be its separation of the evidence based assessment of welfare performance provided through modules 1 - 3 from production philosophies that at times make claims about welfare. Examples of production philosophies that are at risk of being conflated with welfare performance in the consumer's mind include organic, GM free, free of added growth hormones, some animal confinement practices, harvesting practices, growth paths (e.g. slow grown) and so on. A common feature of these is that they are Input (resource or management based), not animal based measures. Thus they may indicate (lowered) risk, but in themselves do not measure good welfare. Modules 1 to 3 enable welfare performance of the enterprise to be assessed independently of the production philosophies brought to livestock management. The latter, nonetheless, have an important role in the market place in servicing the diverse ethical aspirations of consumers.

8. Recommendations

1. Parts 1 and 2 of this review be submitted for peer review to 2 or 3 eminent animal welfare scientists such as Kevin Stafford, Andrew Fisher, and Harry Blokhuis.
2. Part 2 of the review be submitted to key industry stakeholders for review and comment.
3. A Unified Field Index for the assessment of animal welfare on-farm and a process for implementation has been developed for consideration by Australia's livestock industries.

9. References

Bartussek, H 1999. A review of the animal needs index (ANI) for the assessment of animals' well-being in the housing systems for Austrian proprietary products and legislation. Livestock Production Science 61 179–192

Binder, CR, Feola, G, Steinberger, JK 2010. Considering the normative, systemic and procedural dimensions in indicator-based sustainability assessments in agriculture. Environmental Impact Assessment Review 30, 71-81.

EFSA Panel on Animal Health and Welfare, 2012. Guidance on Risk Assessment for Animal Welfare. EFSA Journal 10, 2513-2543.

Croney, CC, Anthony, R 2010. Engaging science in a climate of values: Tools for animal scientists tasked with addressing ethical problems. Journal of Animal Science 88, E75-E81.

Dawkins MS, Donnelly CA, Jones TA 2004. Chicken welfare is influenced more by housing conditions than by stocking density. Nature 427, 342-344.

EFSA Panel on Animal Health and Welfare, 2012. Statement on the use of animal-based measures to assess the welfare of animals. EFSA Journal 10, 2767-2796.

Fabrega E, J Coma, J Tibau, X Manteca and A Velarde 2007 Evaluation of parameters for monitoring welfare during transport and lairage at the abattoir in pigs. Animal Welfare 16 (2): 201-204

Fahmy, M.H., Bernard, C., 1971. Causes of mortality in Yorkshire pigs from birth to 20 weeks of age. Can. J. Anim. Sci. 51, 351-359.

Gonzalez, L.A., Schwartzkopf-Genswein, K.S., Bryan, M., Silasi, R., Brown, F., 2012. Benchmarking study of industry practices during commercial long haul transport of cattle in Alberta, Canada. Journal of Animal Science 90, 3606-3617.

Grandin T (2000) Effect of animal welfare audits of slaughter plants by a major fast food company on cattle handling and stunning practices. JAVMA 216: 848-851.

Grandin T. 2001 Cattle vocalisations are associated with handling and equipment problems at beef slaughter plants Applied Animal Behaviour Science 71(3): 191-201.

Grandin T. 2010 Auditing animal welfare at slaughter points Meat Science 86: 56-65.

Grandin T. 2012 Developing measures to audit welfare of cattle and pigs at slaughter. Animal Welfare 21: 351-356.

Hatcher, S., Eppleston, J., Graham, R.P., McDonald, J., Schlunke, S., Watt, B., Thornberry, K.J., 2008. Higher weaning weight improves post weaning growth and survival in young Merino sheep. Australian Journal of Experimental Agriculture 48, 966-973.

Janhunen, M, Kause, A, Jarvisalo, O 2012. Costs of being extreme - Do body size deviations from population or sire means decrease vitality in rainbow trout? Aquaculture 370, 123-129.

Johnsen PF, Johannesson T, Sandoe, P 2001 Assessment of Farm Animal Welfare at Herd Level; Many Goals, Many Methods. Acta Agriculturae Scandinavica, Animal Science 51:S30, 26-33.

Jones TA, Donnelly, CA, Dawkins MS (2005). Environmental and management factors affecting the welfare of chickens on commercial farms in the United Kingdom and Denmark stocked at five densities. Poultry Science 84, 1–11.

Kanis E, De Greef, KH, Hiemstra, A, van Arendonk JAM 2005. Breeding for societally important traits in pigs. Journal of Animal Science 83:948–957.

Keeling, LJ, Immink, V, Hubbard, C., Garrod, G., Edwards, S.A., Ingenbleek, P., 2012a. Designing animal welfare policies and monitoring progress. Animal Welfare 21, 95-105.

Keeling, LJ, Wallenbeck, A, Larsen, A, Holmgren, N 2012b. Scoring tail damage in pigs: an evaluation based on recordings at Swedish slaughterhouses. Acta Veterinaria Scandinavica 54.

Koeck, A, Miglior, F, Kelton, DF, Schenkel, FS 2012. SHORT COMMUNICATION: Genetic association of body condition score with disease resistance in first lactation Canadian Holsteins. Can. J. Anim. Sci. 92, 285-289.

Lacy, J, 2011. Cropcheck: Farmer benchmarking participatory model to improve productivity. Agricultural Systems 104, 562-571.

Lawrence AB, Conington J, Simm G 2004. Breeding and animal welfare: practical and theoretical advantages of multi trait selection Animal Welfare 13 (supp 1): 191-196.

Leeb, C 2011. The Concept of Animal Welfare at the Interface between Producers and Scientists: The Example of Organic Pig Farming. Acta Biotheoretica 59, 173-183.

Lidauer, M, Mäntysaari, EA, Strandén, I, Pösö, J 2000. Multiple-trait random regression test-day model for all lactations. Interbull bulletin, (25), 81.

Main, DC, Webster J 2011. Assessment, Implementation and Promotion of Farm Animal Welfare In: Management and Welfare of Farm Animals Ed. 5, UFAW Farm Handbook. Wiley Blackwell, Chichester, UK. pp. 564-584.

Napolitano, F, De Rosa, G, Braghieri, A, Grasso, F, Bordi, A, Wemelsfelder F 2012. Qualitative Behavioural Assessment of dairy buffaloes. Applied Animal Behavioural Science 141, 91–100.

Rutherford, KMD, Donald, RD, Lawrence, AB, Wemelsfelder F 2012. Qualitative Behavioural Assessment of emotionality in pigs. Applied Animal Behaviour Science 139: 218– 224.

von Borell E 2000. Assessment of pig housing based on the HACCP concept - critical control points for welfare, health and management. In: Blokhuis, H.J., Ekkel, E.D. and Wechsler, B. (editors): Improving Health and Welfare in Animal Production, EAAP publication No. 102, Wageningen Pers, The Netherlands, pp 75-80.

von Keyserlingk, M.A.G., Barrientos, A., Ito, K., Galo, E., Weary, D.M., 2012. Benchmarking cow comfort on North American freestall dairies: Lameness, leg injuries, lying time, facility design, and management for high-producing Holstein dairy cows. J Dairy Sci 95, 7399-7408.

Appendix 1 - The Unified Field Index for Assessment of Livestock Welfare Performance

Module				Measures			
	Domain	**Class**	**Subclass**	**Risk Assessment**	**Risk Monitoring**	**Self Audit**	**External audit**
1. Animal	Behaviours	Abnormal					
		Social	Agonistic				
			Affiliative				
		Self-care					
	Health	Mortality	Number/rate				
			Causes				
			Found dead				
			Euthanased				
		Morbidity	Numbers/ rates				
			Causes				
		Current status	Skeletal				
			Soft tissue				
			Coat/pelage				
			Demeanor				
	Affect	Does demeanor belong here?					
	Production	Targets	means				
			variances				
	Reproductive performance						
	Holistic attributes						

2. Resources	Feed	Food on offer					
		Quality					
	Water						
	Climate						
	Social	Density					
		Group structure					
	Comfort	bedding					
		Housing/ infrastructure					
	Hygiene						
3. Management	Skills	Training / experience					
		Attitudes					
	Husbandry practices	Methods					
		Analgesia					
		Euthanasia methods					
	Genetics management	Suitability of genetics for environment					
		Breeding objectives for welfare traits,	EBV based				
			Non-quantitative selection criteria				

		disease resistance, temperament, conformation					
		Culling criteria					
	Records	Training					
		Husbandry					
		Vaccinations					
		Medications					
		Farm chemicals					
		Production					
	Review and action protocols						
4. Optional module: Other ethical criteria	Labour	Compliance with ILO standards, UN Human Rights Charter (imported products) and local regulations					
		WOOFer					
	Capital	"ethical" sources					
	Carbon	CO2 equivalents emissions/ capture					
	Environmental sustainability	Water footprint					

		Habitat management					
	Production philosophies	Organic					
		Genetic practices					
		Hormone use					
		Confinement practices					
		Wild caught					
		Slow grown					
	Religious practices	Halal					
		Shechita					
	Other						

Appendix 2 – AssureWel Dairy Cattle Assessment Protocol

Dairy Cattle
Assessment protocol

AssureWel
Advancing Animal Welfare Assurance

Guidance on sampling:

Individual measures	
1a. Mobility – individual scoring **2. Body condition** **3. Hair loss, lesions and swellings** **4. Dirtiness**	• Assessed on **20 cows** from the main milking herd, sampled randomly from all groups by assessor. Lying cows need not be included if it might risk their welfare. • The same 20 cows should be assessed for all 4 individual measures • **3** or more cows out of the 20 to be assessed jointly with the stockperson (record assessor's score only) • **Body condition** to be assessed on an additional **5 dry cows** if managed as a separate group

Herd measures	
1b. Mobility – lameness management **5. Lying comfort** **6. Broken tails** **7. Response to stockperson** **8. Cows needing further care**	• Assessed across **all** animals on farm, this includes milking cows, dry cows, in-calf heifers, calves, any animals in hospital pens and animals due to be culled or leave the farm e.g. cull cows or male calves • **Lying comfort** assessed on animals seen in cubicle housing only

Records measures	
1c. Mobility – verifying self-assessment **9. Mastitis** **10. Heifer and cow survivability**	Assessed across **all** animals on farm, this includes milking cows, dry cows, in-calf heifers, calves, any animals in hospital pens and animals due to be culled or leave the farm e.g. cull cows or male calves.

1. Mobility

a) Individual scoring — **Individual measure**

Sample: **20 cows** selected at random (**3** or more assessed jointly with stockperson)

Assess using the DairyCo scoring method. Observe cows, ideally on a hard (i.e. concrete) non-slip surface. Monitor each cow individually allowing them to make between 6-10 uninterrupted strides. Watch the cow from the side and the rear.

Scoring:

0/1 = Good/Imperfect mobility
Walks with even weight bearing and rhythm on all four feet, with a flat back; long fluid strides possible; or steps uneven (rhythm or weight bearing) or strides shortened; affected limb/s not immediately identifiable

2 = Impaired mobility
Uneven weight bearing on a limb that is immediately identifiable and/or obviously shortened stride (usually with an arch to the centre of the back)

3 = Severely impaired mobility
Unable to walk as fast as a brisk human pace (cannot keep up with the healthy herd) and signs of impaired mobility (score 2)

b) Lameness management — **Herd measure**

Assess and comment on the management of any score 3 cows seen during the visit, including any in a hospital pen. If no score 3 cows are identified ask about the management of the last score 3 cow on the farm.

c) Verifying self-assessment — **Records**

- Check evidence of training in DairyCo mobility scoring
- Verify and comment on the farm's self-assessment of lameness by checking mobility scoring sheets, if they are being completed. Check frequency and scope (e.g. whole herd) of mobility scoring.

2. Body condition **Individual measure**

Sample: **20 cows** selected at random (**3** or more assessed jointly with stockperson)
If the dry cows are kept separately, assess an **additional 5 dry cows** selected at random.

Visually assess cattle based on the Defra condition scoring method, viewing the animal from behind and from the side, the tail head and loin area. Manual assessment can help distinguish borderline scores but may not be necessary for defining the following categories:

Scoring: **Thin = Defra score 1 to less than 2**
Score 1: Tail head – deep cavity with no fatty tissue under skin. Skin fairly supple but coat condition often rough.
Loin – spine prominent and horizontal processes sharp.

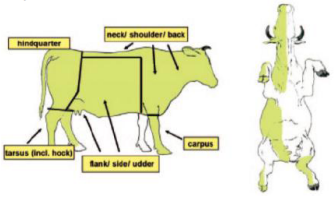
Moderate – Good = Defra score 2 or 3 to less than 4
Score 2: Tail head – shallow cavity but pin bones prominent; some fat under skin, skin supple.
Loin – horizontal processes can be identified individually with ends rounded.
Score 3: Tail head – fat cover over whole area and skin smooth but pelvis may be seen.
Loin – end of individual horizontal process cannot be seen; only slight depression in loin.

Fat = Defra score 4 to 5
Score 4/5: Tail head – completely filled or buried and folds and patches of fat evident.
Loin – cannot see horizontal processes and completely rounded appearance (a slight loin depression may still be seen)

3. Hair loss, lesions and swellings **Individual measure**

Sample: **20 cows** selected at random (**3** or more assessed jointly with stockperson)

Visually assess the following regions of one (randomly selected) side of the animal, from a distance not exceeding 2m (see picture):
a. Head, neck, shoulder, back;
b. Flank, side, udder
c. Hindquarter;
d. Front leg (carpus)
e. Hind leg and hock (include whole of tarsus to coronary band)



© Welfare Quality

Scoring: **0 = No/slight skin damage**
No lesions or hairless patches ≥2cm diameter
1 = Hairless patches
One or more hairless patches (may include scars) ≥2cm diameter
2 = Lesions and/or swelling
One or more lesions (areas of skin damage i.e. wound or scab) or swellings ≥ 2cm diameter.
(Score as a lesion/swelling even if accompanied by a hairless patch. Do not include scars)
NB – swollen hocks = a thickening of the joint such that the usual joint anatomy becomes poorly defined or obscured

4. Dirtiness **Individual measure**

Sample: **20 cows** selected at random (**3** or more assessed jointly with stockperson)

Visually assess the following regions of one (randomly selected) side of the animal and behind:
a) lower hind legs (above the coronary band), including the hock
b) hind quarters – upper hind leg, flank and rear view, including tail (excluding udder)
c) udder

Scoring: **0 = Clean**
No dirt or only minor splashing present (except teats which must have no minor splashing present)
1 = Dirty
An area of dirtiness (i.e. layer or plaques of fresh or dried dirt) amounting to at least palm size (10x15cm). Or, any dirt (including minor splashing) on or around the teats. Do not score stained hair.
2 = Very dirty
An area of dirtiness (i.e. layer or plaques of dirt) amounting to at least forearm length (40cm) in any dimension

5. Lying comfort	Herd measure (cubicle housed animals only)
Whilst assessing the herd, record the number of animals which are not lying correctly, i.e. lying partly (the edge of the cubicle is in contact with the hindquarters or udder) or completely outside the cubicle, or with any other lying difficulty such as dog-sitting or lying backwards. Do not include cows whose heads or bodies are across other cubicles.	

6. Broken tails	Herd measure
Whilst assessing the herd, record the number of animals that show evidence of a broken tail, including tails that are bent, short or injured. Investigate and record possible causes of any broken tails observed.	

7. Response of cattle to stockperson	Herd measure
Check whether the person present for the assessment is the regular stockperson. Throughout the visit, observe the response of the cattle to the stockperson as they approach and interact with the cattle. As far as possible assess response to the stockperson alone, rather than the assessor. Score and comment. **Scoring: 0 = Sociable (to the stockperson)** **1 = Indifferent** **2 = Cautious**	

8. Cows needing further care	Herd measure
Assess the whole herd and record number of any sick or injured cows that would benefit from further intervention. Further interventions could include further treatment, hospitalisation (i.e. removal from the main herd) or culling. Assess animals across the herd including the milking herd, dry cows, in-calf heifers, calves, hospital pens and animals that are due to leave the farm.	

9. Mastitis	Records
Check farm records and record the number of individual clinical cases of mastitis that received treatment of any kind treated in the past 12 months.	

10. Heifer and cow survivability – under development to be implemented soon	Records