

2. Assessing the Welfare of Farm Animals – A Review

Development and Implementation of a Unified field Index (UFI)

February 2013

Ian Colditz¹, Drewe Ferguson¹, Teresa Collins², Lindsay Matthews³ and Paul Hemsworth⁴.

¹ CSIRO Animal, Food and Health Sciences

² Murdoch University

³ Lindsay Matthews & Associates Research International

⁴ Animal Welfare Science Centre



Progress Report - APL Project No 2011/1036.421 - Identify and integrate measures of animal welfare that meet the needs of animals and society.

Disclaimer: APL shall not be responsible in any manner whatsoever to any person who relies, in whole or in part, on the contents of this report unless authorised in writing by the Chief Executive Officer of APL.



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 through 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 levies 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 through 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 (eg. 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 1 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 1, whilst animal-based measures generally provide more accurate indications of the animal’s state and have more universal utility (ie. 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 (eg. 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 (eg. 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 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 intra-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 its 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 production 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 spokesperson 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).

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 exceeded but not continually improved upon. Welfare performance is measured through benchmarking

of performance across enterprises within an industry. Thus the program measures what is achieved and what can be achieved in real world 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.

Welfare Quality Assessment: Dairy Cattle		
http://www.welfarequality.net/publicfiles/36059_25646376170_200705090907523_2244_Proceedings_2nd_WQ_Stakeholder_conference_3_4_May_2007.pdf		
	Welfare Criteria	Measures (all on farm)
Good feeding	1 Absence of prolonged hunger	Body condition score (percentage of too fat/too thin animals)
	2 Absence of prolonged thirst	Water supply (number of water bowls, flow rate, cleanliness, functioning of bowls)
Good housing	3 Comfort around resting	Time needed to lie down Percentage of animals colliding with housing equipment during lying down Percentage of animals lying with hindquarter on edge Cleanliness scores (udder, flank and upper legs, lower legs)
	4 Thermal comfort	
	5 Ease of movement	Presence of tethering Access to outdoor loafing area and/or pasture

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.

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.

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.

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 (eg 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.

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 (eg 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 postweaning 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, disease resistance, temperament,	EBV based				
			Non-quantitative selection criteria				

		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 equivalent 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

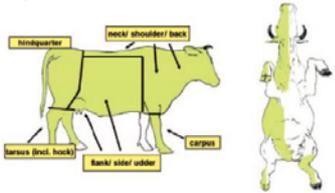
Assessment protocol

Guidance on sampling:

Individual measures	
1a. Mobility – individual scoring 2. Body condition 3. Hair loss, lesions and swellings 4. Dirtiness	<ul style="list-style-type: none"> 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	<ul style="list-style-type: none"> 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):	
	<ul style="list-style-type: none"> 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 $\geq 2\text{cm}$ diameter
	1 = Hairless patches	One or more hairless patches (may include scars) $\geq 2\text{cm}$ diameter
	2 = Lesions and/or swelling	One or more lesions (areas of skin damage i.e. wound or scab) or swellings $\geq 2\text{cm}$ 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:	
	<ul style="list-style-type: none"> 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