Providing cross-species comparisons of animal welfare with a scientific basis

M.B.M. Bracke

Animal Sciences Group, Wageningen University and Research Centre, P.O. Box 65, NL-8200 AB Lelystad, The Netherlands (tel: +31-320-238205; fax: +31-320-238094; e-mail: marc.bracke@wur.nl)

Received 9 November 2005; accepted 10 March 2006

Abstract

Animal welfare issues may involve different species and require decision-makers to compare welfare across species. Up to now applied ethologists have largely ignored questions involving cross-species comparisons. This paper discusses the question whether cross-species comparisons about animal welfare can be provided with a scientific basis, i.e., based on scientific arguments. The arguments pro and contra are reviewed. Conceptually, cross-species comparisons should be possible, but at the explanatory and operational levels substantial problems remain to be resolved. An example is given comparing the welfare of laying hens in battery cages, conventionally housed fattening pigs, conventionally housed broilers and dairy cattle at pasture. Possibly a method could be developed that makes welfare assessments across species more transparent and coherent, and that is based on available scientific information. An outline of such a method is described in this paper.

Additional keywords: farm animal welfare assessment, livestock production, housing systems

Introduction

Animal welfare is a persistent issue in many modern-day societies. To help resolve questions as to what really matters to animals, applied ethologists have been studying the behaviour, stress physiology and pathophysiology of different species of farm animals under a wide range of conditions. Over the years much scientific knowledge about animal welfare has been generated (summarized in various reports and books, e.g. Fraser & Broom 1990; Anon., 1997), even though much remains to be debated (e.g. Kennedy, 1992; Rushen & De Passillé 1992; Stafleu *et al.*, 1996; Bracke *et al.*, 1999a). Despite this 'uncertainty', it is commonly agreed amongst researchers in this field that, at least within species, applied ethologists can compare the different welfare states of animals in different housing systems (Anon., 2001a; Bracke *et al.*, 2002a, b; Rushen, 2003).

Policy makers and consumers, however, must make decisions about animal welfare involving cross-species comparisons (CSCs). At the political level CSC issues are becoming increasingly evident. For example, animal welfare legislation is often formulated in species-specific rules, and it may be questioned why in Europe, for example, calves but not piglets may be taken from their mothers immediately after birth. In addition, a decision on banning mink farming may have to face the question as to whether minks are worse off than, for example, hens in cages. CSCs are not only relevant for political decision-making, also individual consumers may have to make CSCs, e.g. when they decide which animal products they find (ethically) acceptable. If policy makers or consumers are faced with factual questions concerning CSC they may turn to scientists for an answer. Scientists have a common (biological) conceptual framework for welfare assessment and could perhaps answer questions about CSC as purely descriptive questions about what is factually the case with respect to the animals' welfare states. Up to now, however, applied ethologists have largely ignored these questions, possibly because scientists may believe that CSC is beyond their capabilities.

In this paper it is presumed that animals actually have a welfare state that can be assessed by scientists using behavioural and physiological measurements. Building upon these assumptions this paper explores if and how descriptive questions about cross-species comparisons of animal welfare could be answered.

Methods

I try to design a scientific approach to cross-species comparison, especially in relation to farm animal welfare assessment, by presenting a mixture of philosophical, ethical, political and scientific arguments. The arguments concerning cross-species comparisons (CSCs) are classified into arguments at the conceptual, explanatory and operational level (after Stafleu *et al.*, 1996). Conceptual arguments deal with philosophical issues around the definition of terms and basic assumptions about metaphysics/ontology (what is the case) and epistemology (what can be known). Explanatory arguments relate to the scientific (biological) framework. Operational arguments concern the practice of performing CSC. At each level the arguments will be presented and discussed in order to evaluate whether and how scientists could make cross-species comparisons.

At the end of this paper an example will be presented to identify an outline of a method that could be adopted to perform CSCs. It involves a comparison of four common husbandry systems for livestock production, using a list of needs derived from Bracke *et al.*, (1999b). The production systems are: (1) laying hens in battery cages, (2) conventionally housed fattening pigs (small groups of pigs on half-slatted concrete floors without straw), (3) conventionally housed broilers (very large groups on litter substrate), and (4) dairy cattle at pasture (with milking twice a day in a milking parlour). Each system is ranked and weighted 'subjectively', i.e., based on the author's opinion, on a 4-point scale from 1 (need state probably considerably frustrated) to 4 (need state probably satisfied). To arrive at a method to perform CSC, subjective judgments can be replaced progressively by scientific facts from which conclusions are drawn using rules of logic such as a Wilcoxon signed ranks test (Anon., 2001b) to 'test' the hypothesis whether the systems differ in the 'amount' of need frustration and satisfaction.

Conceptual level

At the conceptual level the following arguments pro and contra CSC can be identified. CSC concerns the assessment of the welfare states of animals from different species, i.e., it involves the concepts 'species', 'welfare' and 'assessment'. In this paper welfare is defined as the quality of life as perceived by the animals, i.e., the amount of suffering and happiness experienced by the animals (Anon., 2001a). However, the arguments presented here also hold for other definitions of animal welfare, including definitions in terms of biological functioning (e.g. Broom, 1996).

An assessment of the welfare state of an individual animal entails using some sort of scale, e.g. an ordinal scale including 'good welfare', 'average welfare' and 'poor welfare' (Broom, 1996). Such an assessment involves a reasoning process based on quantitative and qualitative facts and inferences leading to a conclusion where the welfare state is expressed on the scale. For instance, we would like to be able to say 'this animal's welfare level is 8 on a scale from 0 to 10' (Bracke *et al.*, 2002a). Both, welfare experts and lay persons are able to give such welfare scores for various livestock production systems (M.B.M. Bracke, unpublished data). Both groups performed this task without much hesitation, thereby (tentatively) suggesting that they used a common scale across species. This is in accordance with common sense where the metaphysical/ontological assumption is that the welfare of different species can be expressed on a single scale. This assumption underlies the (semantic) fact that only one word for 'welfare' is used for both humans and animals, which can also be found in certain pieces of legislation such as the Dutch law for the protection of laboratory animals. This law requires researchers to weight animal suffering against human interests and it also requires them to register the anticipated level of discomfort on a common scale for different species. From these metaphysical/ontological, psychological, linguistic and political perspectives questions can be raised about CSC, which are conceptually legitimate.

At the conceptual level, CSC issues also are in accordance with the scientific perspective. Welfare can be expressed on a (semi-)quantitative scale and a proper assessment requires an interpretation of scientific facts based on scientific reasoning. Since welfare cannot be measured directly (yet), an assessment of animal welfare is always an assessment from a human's point of view (Bekoff *et al.*, 1992). Science can deal with such items, provided the underlying concepts can be defined adequately and a methodology can be specified to operationalize the assessment (Bracke *et al.*, 2002a). An important aspect is that welfare assessment is typically a judgement that involves uncertainty. The available information is rarely, if ever, sufficient to actually prove, statistically or otherwise, that a given level of welfare is present (Bracke *et al.*, 1999a). CSC requires a reasoning process leading to a conclusion (Rushen, 2003). The reasoning process may differ in detail, must follow the rules of logic and rational thinking, and ideally should result in the most probable and best possible assessment given the present state of the art in science. This is true for a welfare assessment within species, but even more so if a comparison is made across species. Although much uncertainty remains, some cases involving CSC nevertheless seem to be clear enough to conclude that in principle CSC is possible. For example, a starving rat almost certainly is experiencing more hunger than a well-fed dog. Similarly, few scientists would deny that tethering systems for individually confined pregnant sows are worse for welfare than keeping herds of cattle in semi-natural enclosures. Most humans are not only (widely believed to be) more intelligent, but their welfare status too is better than that of most farmed domestic fowl. Conversely, some human suffering is substantial and would seem to factually outweigh certain types of animal suffering, thereby providing a *prima facie* ethical justification for animal experimentation to the scientists performing the experiments. These clear cases suggest that CSC is within the scientific domain and also that CSC is in principle possible from a scientific point of view. So the challenge for a methodology to perform CSC in less clear cases is to reduce the degree of uncertainty by providing (at the explanatory and operational levels) empirical data, decision rules and a conceptual framework to make the entire reasoning process, i.e., all steps between premises and conclusion, more reliable and transparent. As long as the considerations underlying the assessment remain implicit, the judgement cannot be distinguished from anthropomorphic, 'emotional' and subjective judgements. But, if the reasoning steps are made explicit and if uncertainties are exposed to critical evaluation, the role of the evaluating subject is reduced and apparent personal judgements can become intersubjective, even 'objective' in the end. When science has established such an objective methodology to assess the welfare status of animals within and between different species as a factual issue, decision makers can then use the results of such factual welfare assessment as input in the ethical and political process of deciding what levels of welfare are acceptable, both individually and for society as a whole.

Explanatory level

At the explanatory level a comparison of welfare across species must have an explanation within some conceptual framework where it plays a functional role. Gosling (2001) provides the following illustration.

'I tell you a black mamba is trapped in the room next door, and you ask me whether it is aggressive. If I adopt a within-species framework, I may respond, "No, it is very unaggressive," because it has attacked only two persons in the last hour, well below the norm for this species of snake. If on the other hand, I adopt a cross-species framework, I may respond, "Yes, of course it's aggressive," because it is a black mamba, a highly aggressive species of snake. Thus, the framework I adopt will determine whether I answer yes or no to your question about the black mamba's level of aggression.' (Gosling, 2001; pp. 58–59). Adopting a CSC framework 'works' in daily life as well as in science.

In biology the law of parsimony is respected. It states that the simplest explanations are to be preferred (Martin & Bateson, 1990). It has, for example, been used to deny that animals have feelings and emotions (e.g. Kennedy, 1992). However, when it is accepted that animals have a welfare state, as is done in this paper, then the law of parsimony requires that similar welfare measurements in different species should be interpreted similarly, unless there are good (empirical) reasons to do otherwise. So as a formal requirement the law of parsimony supports CSC.

Further support for CSC derives from the Darwinian idea that species have evolved from common ancestors, as evidenced not only by anatomical and embryological characteristics (e.g. Slack, 2000) but also by behavioural and stress-physiological response patterns (e.g. Denver, 1997; Conlon, 2000). Several causal models of behaviour and stress have been proposed that apply across species (Bakker & Main, 1980; Ng, 1995). In the field of farm animal welfare assessment a common conceptual framework among welfare scientists has been formulated (Anon., 2001a). In this paper I shall illustrate the complexities involved in CSC in relation to a simplified version of that framework, the Istwert–Sollwert (I–S) model of motivation (Wiepkema, 1987).

In the course of evolution, animals have acquired cognitive-emotional systems – which I shall refer to as 'needs' – to deal with (predict and control) variable circumstances. According to the I–S model, welfare problems arise when animals fail to adapt, i.e., when they structurally fail to match the actual state of their environment (as they perceived it, 'Istwert') with the desired state ('Sollwert') using their behavioural or physiological responses. Istwert-perceptions depend on the animals' sensory system. Sollwerte are the result of both the (ontogenetic) life history of the animals, but especially also of the (phylogenetic) evolutionary history of the species in its ecological niche, its environment of evolutionary adaptation. In order to assess the welfare status of a herd of farm animals of the same species and age group we must somehow totalize (add up, integrate) the amount of satisfaction and frustration of all the different needs, where the 'amount' is a function of the intensity, duration and incidence of the need states of the animals in the group.

Compared with an assessment of welfare of animals of the same species in different housing systems, an assessment across species poses some additional, but at the theoretical level not insurmountable complications. A CSC must take into account that animals of different species may be exposed to different environmental stimuli. They may have different sensory and cognitive systems (and concomitant different Istwert-perceptions), have different Sollwerte and show different behavioural and physiological responses as a result of the different phylogenetic and ontogenetic backgrounds. For example, the different evolutionary histories between specialist and generalist species have resulted in differences in the degree to which they perceive and respond to challenges.

Cross-species comparisons can be performed if apparent differences can be reduced to underlying functional commonalties. Emotions have been linked to only a few brain structures (the 'limbic system') that show certain similarities between individuals and species (Panksepp, 1998), but the way the emotions are triggered and translated into behavioural and physiological responses may differ widely between individuals and species. For example, although a bird may fly, a fish swim and a deer run from danger, these different behaviours may serve the same function, i.e., escaping from danger (Plutchik, 1980). Similarly, food acquisition in pigs involves rooting, in poultry it involves scratching and pecking, and in cattle grazing and rumination.

Even though these behaviours differ qualitatively from one another, the degree to which species-specific foraging behaviours can be performed can be assessed if the environmental conditions are known. A dairy cow at pasture is able to perform foraging behaviour whereas a pig housed on a concrete floor and a hen in a battery cage are not. The importance of performing a certain type of behaviour depends on the (biological significance of the) behavioural and physiological consequences. Grazing is less important for cattle than rooting is for pigs when – other things being equal – thwarted pigs develop abnormal oral behaviours, whereas cattle do not. On the other hand, ruminating is more important for cattle than rooting is for pigs since the inability to ruminate is life-threatening for cattle whereas such risks are not involved for pigs. It is this kind of reasoning that allows drawing conclusions about the importance of environmental conditions. Individual housing may be as important for a solitary species as group housing is for a gregarious species when similar consequences of 'deprivation' are observed. The same concrete floor may pose a different problem for pigs (inability to root, leading to tail biting) than for cattle (uncomfortable resting, leading to lameness).

In all these cases it is the animal's (welfare) performance in terms of its behavioural and physiological responses, that ultimately reflects its level of welfare.

To study treatment effects, applied ethologists normally use groups of animals of the same species, keeping experimental conditions as constant as possible. Normally recognized variables include environmental conditions such as climate and feed, and animal-based conditions such as weight, sex and age. However, some of these within-species variables, e.g. those between different age-groups, certain breeds of poultry (broilers versus laying hens), may be larger than some between-species variables such as between horses and donkeys, between certain breeds of domestic fowl and quail, or between cattle and buffalo. For example, Duncan (1979) found profound differences between a 'flighty' and a 'docile' strain of hens. The 'flighty' hens showed extreme panic and a brief elevated heart rate to visual stimulation, whereas 'docile' strains showed less extreme behaviour but a prolonged heart-rate elevation. In addition, Korte *et al.* (1997) and Van Hierden *et al.* (2002) found substantial behavioural and stress-physiological differences among different lines of White Leghorns, with a high and low propensity to perform feather pecking. If such comparisons are allowed, it seems arbitrary not to allow at least some types of cross-species comparisons.

In applied ethology, frequencies and time budgets e.g. of alarm responses and agonistic interactions in a group of animals of the same species are commonly used as welfare measures. It seems to be assumed that the intensities of the measures co-vary with the frequencies and that each incidence counts for one regardless of the distribution over the animals in the group. For example, n = 5 bites, regardless of whether one animal is bitten 5 times or 5 animals are bitten once. If such assumptions are warrant-ed within species, for practical reasons or otherwise, similar assumptions can be formulated for making comparisons across species. If the results are interpreted within a biological framework, the most common welfare-related measures such as preference tests, measures of cortisol and heart rate variability, human-approach fear tests, novel object tests, fear-potentiated startle tests (e.g. Davis *et al.*, 1993), elevated plus maze tests, novel environment tests and open field tests can be used to make comparisons across species. The point here is that at the explanatory level and despite all remain-ing uncertainties, the facts about the biology of different animal species can also be interpreted in a biological framework such as the Istwert–Sollwert model so as to make CSC about animal welfare explicit and transparent.

Operational level

At the operational level the question arises as to how to assess welfare in different species.

A statement such as 'wild animals differ in many respects from domesticated species' (Deag, 1996) suggests that in some cases comparisons across species are already operational. One such difference concerns a reduction in brain weight (Herre & Röhrs, 1990). Operationalization involves making valid interpretations of selected welfare measures. Ideally, these welfare measures should apply across species and allow unambiguous welfare interpretations. Welfare measures that differ across species may be used provided they allow a common interpretation. For example, the wallowing of the pig and the moistening of the rat's fur could both indicate a certain level of overheating. In such cases we look for the common denominator, like rooting in pigs and scratching in poultry both being referred to as 'foraging behaviour', and tongue-rolling in veal calves and stereotypic pacing in broiler breeders both being referred to as 'abnormal behaviour'.

If common denominators are used, 'species' may become a factor, an independent variable in the design of a study. Such analyses have already been reported, be it sporadically. Bashaw *et al.* (2001) showed that sub-species was a factor in explaining variation in stereotypic behaviour in different ungulates (see also Leenaars *et al.*, 1998 and Marx *et al.*, 1999). In fields like animal cognition and (psycho)medical research CSCs are more common. Laboratory-animal models are being used to extrapolate findings to human beings (e.g. Ford & Hornby, 1996; Smith *et al.*, 2001). In the field of psychopharmacology, animal models are being used to test psychoactive drugs such as anti-depressants and anxiolytic drugs. Such comparisons are based on similarities in neurophysiological architecture (homology between species). Operationally such extrapolations often work and they presuppose CSC.

In literature occasionally statements can be found that presume CSC on animal welfare. Cross-species comparisons are sometimes made in terms of a trait common to both species; for example, firemouth fish are more fearful than goldfish (Shaklee, 1963), and lion-tailed macaques are more curious than cynomolgus macaques (Clarke & Lindburg, 1993). Other cross-species comparisons concern the existence or non-existence of a trait in different species; for example, a dimension of curiosity appears to be present in meerkats but absent in Cuban ground iguanas (Glickman & Sroges, 1966).' (Gosling, 2001; p. 58). Morton & Griffith (1985) published a paper on the assessment of 'pain, distress and discomfort' in different species of (laboratory) animals. Stern (2001; p. 99) compared the behaviour of domestic fowl, turkeys, ducks and geese, and identified differences in various needs. These included the need for food (geese are more 'vegetarian'), the need for water (to swim in), movement, safety (the need for cover), thermal comfort, social contact, the need for nesting (broodiness). Bartussek (1986) designed a welfare index that applied to several species of farm animals such as cattle, pigs, poultry and rabbits. In some studies CSCs are being made, e.g. comparing different euthanasia methods (Lambooy, 1985; Iwarsson & Rehbinder, 1993) and comparing Hypothalamic-Pituitary-Adrenal (HPS) Axis responsiveness in different species of cats (Carlstead *et al.*, 1992). Muenker *et al.* (2000) studied the sensitiveness of sea fish on being hauled on board and found that 'among demersal fishes the species without swimbladder and flat fishes were clearly more resistant to mechanical stress. On the contrary, pelagic fish species were generally less robust.' Such differences in physical damage and mortality rates are reliable indicators of reduced welfare in these different species of fish. These examples show that in some respects relevant to welfare assessment CSCs have already been performed at the operational level. To our knowledge no experiments have been performed that directly compared the welfare of different farm animals. To support decision-making in this area, the challenge remains to design studies like comparing the behaviour and (stress) physiology of laying hens and minks in standard and in enriched cages, and more generally, comparing the welfare status of different species of farm animals for the purpose of labeling for animal welfare.

An example of cross-species comparison

In a relatively recent paper (Anon., 2001a), scientists formulated a common conceptual framework for overall welfare assessment by ranking design criteria, performance criteria and the main housing systems for different species and age groups of farm animals. In this paper (Anon., 2001a) the question remained whether it were possible to take such prioritization one step further, i.e., across species. Here the outline of a methodology will be presented by way of a tentative example illustrating how CSC could be performed in a transparent and systematic way. In this example four livestock production systems for three different species of farm animals are compared: (1) battery cages for laying hens, (2) conventional housing for fattening pigs, (3) conventional housing for broilers, and (4) dairy cattle at pasture. If one seriously wonders what would be the welfare states of the animals in the different systems, then one realizes that the factor 'species' appears to be just an additional factor in an already rather complex task. Note that layers and broilers belong to the same species, but this does not seem to make the comparison substantially more feasible in this pair of systems compared with the other pairs. In fact the differences between different types of animals belonging to one species but housed in different systems, such as layers in cages versus conventionally housed biologically farmed broilers, appear to be larger than the differences between different species in similar systems such as cows and sows grazing on pasture. The outlines of the welfare assessment task within one category of animals (pregnant sows) have been described and operationalized (Bracke *et al.*, 2002a). Welfare assessment requires an assessment of the different welfare needs of the animals (needs derived from Bracke *et al.*, 1999b). For each need, each production system may be ranked relative to the other systems and weighted, e.g. on a 4-point scale from 1 (need state probably considerably frustrated) to 4 (need state probably satisfied).

The results of this evaluation process are presented in Table 1. In this table the four housing systems (layers in cages, fattening pigs, broilers and dairy cattle at pasture) have been scored for each of 18 needs, wherever this was applicable. Table 1 does not pretend to be factually correct. It is intended to illustrate how CSC could be performed. The table is based on the subjective opinions of the author of this paper. However, it is possible to make these opinions explicit, to identify scientific facts underlying these opinions and to make the reasoning process transparent and verifiable. For each (welfare) need Table 1 gives an assessment in terms of the expected discrepancy between the animal's provisions (Istwerte) and its setpoints (Sollwerte). The Sollwerte are derived from the animal's natural behaviour (e.g. social animals have a natural tendency to live in social groups), from its 'demands' and from knowledge about negative welfare symptoms (such as frustration, aggression, stress, abnormal behaviour, reduced fitness). The production systems (e.g. battery cages in the case of the laying hens) define the Istwerte. The rank numbers in the table are intended to represent the animal's comparisons of Istwerte and Sollwerte. It is presumed that the degree of mismatch, i.e., the degree to which each need is satisfied or frustrated, is related to the intensity, duration and incidence of the various positive and negative welfare performance criteria. Clearly, such an assessment cannot be performed without data from empirical research providing both accurate descriptions of the production systems as well as information about the welfare (performance) consequences of keeping animals in these systems (compared with keeping them in more natural or otherwise welfare-friendly conditions).

For example, the need for food is considered to be adequately satisfied in each of the four production systems (ranking = 4), whereas the satisfaction of the need to forage is assessed as 1, 2, 3 and 3.5 for hens, pigs, broilers and cows, respectively. It is based on the considerations about the ability to perform foraging behaviour (depending on space, floor type and the presence of substrates) and the known relationship with abnormal behaviours (such as feather pecking and tail biting). The ranking may be disputed, however, and further development of the method of CSC should include specifying the arguments used to assign the rank-numbers in a more detailed and methodic way. For example, the assessment of the need for safety (fear) may use the following empirical observation: when the farmer enters the barn in a standardized way, different farm animals respond differently. Laying hens and broilers may show hysteria, pile up in panic while killing the birds underneath. Fattening pigs may jump to their feet, briefly give alarm vocalizations and then show inquisitive behaviour. Fattening bulls may hardly respond. From a biological perspective these behavioural differences probably represent different fear levels in the different species in response to the same stimulus (e.g. a fox is a threat to poultry, but not to cattle). This empirical knowledge is relevant to assess the animals' perception of safety, but other information must be taken into account in order to determine the biological significance of the phenomena. For example, the lack of overt responding in bulls must be regarded as indifferent behaviour rather than freezing. Conversely, because hens are lighter and have wings, it may be expected that a pile of hysteric bulls would be experiencing

Table 1. Tentative welfare comparison of four conventional livestock production systems, using a ranking (scale 1–4; 1 = worst; 4 = best) within species for each welfare need. The needs are hierarchically organized, with first-order needs in italics; needs after Bracke *et al.*, 1999b.

Need	Laying hens	Fattening pigs	Broilers	Dairy cows at pasture
Ingestion				
Food	4 *ad libitum*	4 *ad libitum*	4 *ad libitum*	4 *ad libitum*
Water	4 clean nipple, easy to reach	4 clean nipple, easy to reach	4 clean nipple, easy to reach	4 clean trough in pasture
Rest	2 no perch; wire floor	2 no nest; hard surface	3 no perch; litter floor	4 proper surface, pasture
Social contact	3 ca. 4–5 animals per group	3 ca. 8–12 animals per group	2 > 1000 animals per group	3 ca. 40 animals per group
Reproduction				
Sexual contact	1 deprived	prepubertal	prepubertal	2 contact with herd mates; no bull
Nest building	1 no nest box	not applicable	not applicable	2 limited separation from the herd; calf separated from the cow
Maternal contact	not applicable	not applicable	not applicable	calf removed at birth (full udder)
Movement	1 0.045 m^2 per hen in the cage	2 0.7 m^2 per pig in the pen	3 0.05 m^2 per broiler in large barn	4 500 m^2 per cow at pasture

Table 1 continued.

Need		Laying hens		Fattening pigs		Broilers		Dairy cows at pasture
Exploration								
Explore novelty	1	none	1	none	2	some	2	some (with outdoor and milking parlour access)
Foraging	1	wire floor	2	half-slatted floor	3	litter without food rewards	3.5	pasture, no browsing
Play		not applicable (adult animals)	2	little space, slatted floor	3	more space and litter floor		not applicable (adult animals)
Body care	2	preening possible, little space, no wing flapping	2	no scratch post, no wallow	3	preening and wing flapping possible, some soiling	3	allogrooming possible, no scratch post
Evacuation	3	wire floor, no space	2	limited separation of resting and dunging area	3	some pen soiling, some space	3	incidental manure spread on pasture
Thermoregulation	2	no control	3	little control, can huddle, solid and slatted floor area	2	no control	3	more variable, normally no shelter (may be hot or wet)
Respiration	2	NH_3 levels indoors	2	NH_3 levels indoors	1	feels asphyxiated	4	fresh air
Health								
No injuries (no pain)	2	beak trimmed, some feather pecking	2	tail docked, some tail biting, some lameness	1	substantial lameness	3	some mastitis, some lameness (electric fencing)
No (other) illness	2.5	few	2.5	few	2.5	few	2.5	few
Safety (incl. agression)	3	some	3	some	3	some	4	little agression

more fear than a similar pile of hens. An assessment of the actual level of distress must also take into account differences in external stimulation. Farmers may enter buildings in different ways, e.g. farmers often knock on the door before entering a poultry barn, and aversive stimuli may not only come from humans e.g. in the form of veterinary treatments, but also from conspecifics in the form of aggression. This makes the assessment of a welfare need such as safety a rather complex task. It is even further complicated by the fact that in addition to the need for safety an overall welfare assessment requires assessing other needs as well as a procedure to integrate these need states into an overall judgement, prioritization or score. On the other hand, the impact of small assessment errors on the overall scores may be relatively small, e.g. when small errors randomly vary around the mean. In this paper the assumption for integrating different needs into an overall score is that each need is equally important (see Bracke *et al.*, 2002a for an alternative weighting method). With that assumption the means of the rankings and the Wilcoxon's signed ranks test may be used to examine the relative 'amounts' of need frustration in each of the four livestock production systems. The mean scores are 2.16 ($n = 16$), 2.43 ($n = 15$), 2.63 ($n = 15$) and 3.19 ($n = 16$) for hens, pigs, broilers and cows, respectively. The Wilcoxon signed ranks test shows that cows at pasture are judged to differ significantly ($P < 0.05$) from each of the three other systems, whereas these other systems do not differ significantly from one another ($P > 0.05$).

The example shows how, at least in clear cases such as the comparison between battery hens and cows at pasture, comparisons across species may be performed by a scientist in a methodic way using empirical information, a list of needs and an assumption about the integration of rankings. In this example different need states were used as the common denominators for CSC, but it should be realized that selecting a different list of needs or a differential weighting of the different needs may alter the outcome. Realize also that a critical re-evaluation of each system's rank scores may result in minor modifications, but this will probably not lead to a very different overall ranking of the systems for two reasons. Firstly, dairy cattle repeatedly scored better than laying hens as illustrated with the Wilcoxon statistic and, secondly, the dairy cattle have been scored somewhat conservatively, e.g. cattle with outdoor access on pasture and being milked twice daily could be argued to have considerably more opportunities to explore novelty than broilers, pigs or layers in cages. The challenge for more empirically-oriented applied ethologists is to design measurable and validated parameters that apply across species and have a clear welfare interpretation. The challenge for ethicists and political scientists is to incorporate the outcomes of such assessments into decision-making. For example, a broiler farm will comprise many more animals than a herd of cattle. One may legitimately ask whether ethically and/or politically a cow, a broiler chicken and a human being are to be counted equally. In many other respects more work needs to be done such that CSC in the end involves impartial (factual) judgements made in a procedural way based on a thorough knowledge of the facts. The development of a methodology for CSC would not only help resolve some actual decision-making issues, it could also help making more explicit the assumptions implicit in applied ethology.

Conclusions

Cross-species comparisons (CSC) are becoming increasingly important in political and public decision-making and may pose intriguing challenges to applied ethologists. A compilation of philosophical, ethical, political and scientific arguments leads to the conclusion that in principle a scientific basis for CSC can be constructed. However, at present only widely different housing systems can be distinguished as was shown by way of an example comparing laying hens in cages, fattening pigs and broilers in conventional housing, and dairy cattle at pasture. There are good prospects for further improvements of the method used to make this CSC. This could lead to increasing detection levels based on scientific argumentation and the input of factual knowledge.

Acknowledgements

The author wishes to thank the Dutch Ministry of Agriculture, Nature and Food Quality for funding this work. The valuable comments and contributions of Berry Spruijt, Hans Hopster and other colleagues in Lelystad are gratefully acknowledged.

References

Anonymous, 1997. The Welfare of Intensively Kept Pigs. Scientific Veterinary Committee, EEC, Brussels, 190 pp.

Anonymous, 2001a. Scientists' assessment of the impact of housing and management on animal welfare. *Journal of Applied Animal Welfare Science* 4: 3–52.

Anonymous, 2001b. SPSS® Base 11.0 User's Guide. SPSS, Chicago, Illinois, 574 pp.

Bakker, H.R. & A.R. Main, 1980. Condition, body composition and total body water estimation in the quokka, *Setonix brachyurus* (Macropodidae). *Australian Journal of Zoology* 28: 395–406.

Bartussek, H., 1986. Ergänzung zum Artikel 'Vorschlag für eine Steiermärkische Intensivtierhaltungs-verordnung'. Irdning (1985). *Der Österreichische Freiberufstierartzt* 97: 4–15.

Bashaw, M.J., R.L. Tarou, T.S. Maki & T.L. Maple, 2001. A survey assessment of variables related to stereotypy in captive giraffe and okapi. *Applied Animal Behaviour Science* 73: 235–247.

Bekoff, M., L. Gruen, S.E. Townsend & B.E. Rollin, 1992. Animals in science: Some areas revisited. *Animal Behaviour* 44: 473–484.

Bracke, M.B.M., B.M. Spruijt & J.H.M. Metz, 1999a. Overall animal welfare assessment reviewed. Part 1: Is it possible? *Netherlands Journal of Agricultural Science* 47: 279–291.

Bracke, M.B.M., B.M. Spruijt & J.H.M. Metz, 1999b. Overall welfare reviewed. Part 3. Welfare assessment based on needs and supported by expert opinion. *Netherlands Journal of Agricultural Science* 47: 307–322.

Bracke, M.B.M., B.M. Spruijt, J.H.M. Metz & W.G.P. Schouten, 2002a. Decision support system for overall welfare assessment in pregnant sows. A. Model structure and weighting procedure. *Journal of Animal Science* 8: 1891–1834.

Bracke, M.B.M., J.H.M. Metz, B.M. Spruijt & W.G.P. Schouten, 2002b. Decision support system for overall welfare assessment in pregnant sows. B. Validation by expert opinion. *Journal of Animal*

Science 8: 1835–1845.

Broom, D.M., 1996. Animal welfare defined in terms of attempts to cope with the environment. *Acta Agriculturae Scandinavica Section A, Animal Science*, Supplement 27: 22–28.

Carlstead, K., J.L. Brown, S.L. Monfort, R. Killens & D.E. Wildt, 1992. Urinary monitoring of adrenal responses to psychological stressors in domestic and nondomestic felids. *Zoo Biology* 11: 165–176.

Clark, A.S. & D.G. Lindburg, 1993. Behavioral contrasts between male cynomolgus and lion-tailed macaques. *American Journal of Primatology* 29: 49–59.

Conlon, J.M., 2000. Singular contributions of fish neuroendocrinology to mammalian regulatory peptide research. *Regulatory Peptides* 93: 3–13.

Davis, M., W.A. Falls, S. Campeau & M. Kim, 1993. Fear-potentiated startle: A neural and pharmacological analysis. *Behavioural and Brain Research* 58: 175–198.

Deag, J.M., 1996. Behavioural ecology and the welfare of extensively farmed animals. *Applied Animal Behaviour Science* 49: 9–22.

Denver, R.J., 1997. Environmental stress as a developmental cue: Corticotropin-releasing hormone is a proximate mediator of adaptive phenotypic plasticity in amphibian metamorphosis. *Hormones and Behaviour* 31: 169–179.

Duncan, I.J.H., 1979. Some studies on heart-rate and behaviour in the domestic fowl. *Applied Animal Ethology* 5: 294–295.

Ford, S.L. & S.B. Hornby, 1996. The ewe as a model for post menopausal osteoporosis – a comparison of normal calcium metabolism with man. *Animal Technology* 47: 9–18.

Fraser, A.F. & D.M. Broom, 1990. Farm Animal Behaviour and Welfare (3rd edition). Baillière Tindall, London, 437 pp.

Glickman, S.E. & R.W. Sroges, 1966. Curiosity in zoo animals. *Behaviour* 26: 151–188.

Gosling, S.D., 2001. From mice to men: what can we learn about personality from animal research? *Psychological Bulletin* 127: 45–86.

Herre, W. & M. Röhrs, 1990. Haustiere Zoologisch Gesehen (2nd edition). Fischer, Stuttgart, 412 pp.

Iwarsson, K. & C.A. Rehbinder, 1993. A study of different euthanasia techniques in guinea pigs, rats and mice. Animal response and postmortem findings. *Scandinavian Journal of Laboratory Animal Science* 20: 191–205.

Kennedy, J.S., 1992. The New Anthropomorphism. Cambridge University Press, Cambridge, 194 pp.

Korte, S.M., G. Beuving, W. Ruesink & H.J. Blokhuis, 1997. Plasma catecholamine and corticosterone levels during manual restraint in chicks from a high and low feather pecking line of laying hens. *Physiology & Behavior* 62: 437–442.

Lambooy, E., 1985. Electroanaesthesia or electroimmobilisation of calves, sheep and pigs by the Feenix Stockstill. *Veterinary Quarterly* 7: 120–126.

Leenaars, P.P.A.M., M.A. Koedam, P.W. Wester, V. Baumans, E. Claassen & C.F.M. Hendriksen, 1998. Assessment of side effects induced by injection of different adjuvant/antigen combinations in rabbits and mice. Laboratory animals. *Royal Society of Medicine Press* 32: 387–406.

Martin, P. & P. Bateson, 1990. Measuring Behaviour: An Introductory Guide. Cambridge University Press, Cambridge, 222 pp.

Marx, H., S.T. Sengmuller, R. Hoffmann & A. Stolle, 1999. Stress and quality of trout, catfish and flounder after stunning and slaughter. Untersuchungen zu Stressbelastung und Produktqualitat von Forelle, Wels und Flunder in Zusammenhang mit Betaubung und Schlachtung. *Archiv für Lebensmittelhygiene* 50: 37–40.

Morton, D.B. & P.H.M. Griffiths, 1985. Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. *Veterinary Record* 116: 431–436.

Muenker, W., H. Kuhlmann & J. Oehlenschlaeger, 2000. Sensibilitaet von Seefischen an Bord. Teil 2. Demersale und pelagische Fischarten aus Schleppnetzfaengen in der Nord- und Ostsee. *Informationen für die Fischwirtschaft aus der Fischereiforschung* 47: 97–101.

Ng, Y.K., 1995. Towards welfare biology: evolutionary economics of animal consciousness and suffering. *Biology and Philosophy* 10: 255–285.

Panksepp, J., 1998. Affective Neuroscience. Oxford University Press, New York, 449 pp.

Plutchik, R., 1980. Emotion: a Psychoevolutionary Synthesis. Harper and Row, New York, 440 pp.

Rushen, J., 2003. Changing concepts of farm animal welfare: Bridging the gap between applied and basic research. *Applied Animal Behaviour Science* 81: 199–214.

Rushen, J. & A.M.B. De Passillé, 1992. The scientific assessment of the impact of housing on animal welfare: A critical review. *Canadian Journal of Animal Science* 72: 721–743.

Shaklee, A.B., 1963. Comparative studies of temperament: Fear responses in different species of fish. *Journal of Genetic Psychology* 102: 295–310.

Slack, J.M.W., 2000. The xenopus tadpole: A new organism for regeneration research. *E-Biomed: Journal of Regenerative Medicine* 1: 1–4.

Smith, D., P. Trennery, D. Farningham & J. Klapwijk, 2001. The selection of marmoset monkeys (*Callithrix jacchus*) in pharmaceutical toxicology. *Laboratory Animals London* 35: 117–130.

Stafleu, F.R., F.J. Grommers & J. Vorstenbosch, 1996. Animal welfare: evolution and erosion of a moral concept. *Animal Welfare* 5: 77–84.

Stern, A., 2001. Geflügel Natürlich und Artgerecht Halten. Kosmos, Stuttgart, 104 pp.

Van Hierden, Y.M., S.M. Korte, E.W. Ruesink, C.G. Van Reenen, B. Engel, G.A.H. Korte-Bouws, J.M. Koolhaas & H. Blokhuis, 2002. Adrenocortical reactivity and central serotonin and dopamine turnover in young chicks from a high and low feather-pecking line of laying hens. *Physiology & Behavior* 75: 653–659.

Wiepkema, P.R., 1987. Behavioural aspects of stress. In: P.R. Wiepkema & P.W.M. Van Adrichem (Eds), Biology of Stress in Farm Animals: an Integrative Approach. Seminar in the CEC programme of Coordination Research on Animal Welfare, 17–18 April 1987, Oosterbeek. Nijhoff, Dordrecht, pp. 113–133.