



Australian Government
Department of Agriculture



Validation of On-Farm Pig Welfare Indices

Final Report
APL Project 2013/062

September 2014

Animal Welfare Science Centre
The University of Melbourne
Ellen Jongman / Lauren Hemsworth / Jeremy Skuse
Parkville, VIC 3010

Disclaimer: The opinions, advice and information contained in this publication have not been provided at the request of any person but are offered by Australian Pork Limited (APL) solely for informational purposes. While APL has no reason to believe that the information contained in this publication is inaccurate, APL is unable to guarantee the accuracy of the information and, subject to any terms implied by law which cannot be excluded, accepts no responsibility for loss suffered as a result of any party's reliance on the accuracy or currency of the content of this publication. The information contained in this publication should not be relied upon for any purpose, including as a substitute for professional advice. Nothing within the publication constitutes an express or implied warranty, or representation, with respect to the accuracy or currency of the publication, any future matter or as to the value of or demand for any good.

I. Acknowledgements

This project is supported by funding from Australian Pork Limited and the Department of Agriculture. The authors wish to thank the commercial producer who made their facilities available so that this research could be conducted.

2. Executive Summary

The Australian Pork Industry Quality Assurance Program (APIQ) allows producers to be annually audited (PigCare audit) against a set of standards and performance indicators. This audit does not however provide producers and managers with the opportunity to monitor the welfare status of their pigs over time and/or benchmark the welfare status of animals between farms. Given that the Australian Pork Industry adopts a pro-active approach to animal welfare, a protocol which provided a simple, validated set of pig welfare indicators would both enable the industry to make evident its commitment to pig welfare and demonstrate incremental welfare improvements over time.

The current project was performed to examine the external validity and intra- and inter-observer reliability of three modified animal-based welfare indices (BCS, LS, and IS identified during APL Project 2012/1025) and their modified methodologies recommended for a simpler and more practical on-farm assessment of body condition, lameness and injury in group-housed pigs. Once validated, the modified welfare indices will be able to be included in the practical on-site pig welfare benchmarking tool proposed by the earlier APL project 2012/1025. Thus, the current project sought to:

1. To examine the external validity of the modified animal-based welfare indices (body condition score, injury score, lameness score and fear of humans), identified during APL Project 2012/1025, as on-farm measures of pig welfare in the Australian Pork Industry.
2. To investigate the inter- and intra-observer reliability of the modified animal-based welfare indices included in the on-site pig welfare benchmarking protocol.

The external validity and inter- and intra-observer reliability of the modified animal-based welfare indices, body condition score, injury score and lameness score, were investigated in group-housed sows and grower pigs at a large commercial piggery in Regional Victoria, over a 6-week period in June and July, 2014. Sampling was conducted by four trained observers on 240 group-housed pigs over six 2-day periods (120 sows in weeks 1-3 and 120 grower pigs in weeks 4-6); each observer assessed 20 animals for BCS, LS and IS, on day 1 of the sampling period using both the validated (V) and the modified (M) methodologies, and 40 animals using the modified (M) methodologies on day 2 (consecutive days).

The validity of the modified animal-based welfare indices as on-farm measures of pig welfare was investigated by comparing the assessment scores obtained from the modified methodologies and those from the validated methodologies, in order to determine the level of agreement between the two measures. In addition, the inter- and intra-observer reliability of the animal-based welfare indices (BCS, LS and IS) was investigated using a test-retest assessment (Whitham and Wielebnowski, 2009), which examined the similarity between measures collected on an animal at two different time points (intra-observer) and the agreement between measures taken on an animal by multiple assessors (inter-observer). The validity testing largely indicated a moderate relationship with slight to fair levels of agreement between the validated and the modified animal-based welfare indices. The modified animal-based indices were derived from validated methods of body condition, lameness and injury assessment, based on the scientific literature and practical plausibility. This was achieved by simplifying the scoring scales and reducing the subjectivity of the assessment in order to improve their on-farm practicality. Given that the validated animal-based welfare indices underwent only minor modification, it was anticipated that the modified indices would correlate more highly and have better levels of agreement with the validated measures than the results suggest. One explanation for the projects findings is the homogeneity of the data, resulting from a lack of variation

in the animals sampled; most animals were scored in the same category and no animals were assigned scores indicated a clear welfare issue. Further testing could be performed in populations with greater variation in order to confirm the validity of the modified animal-based indices. Whilst the correlations and the level of agreement between the modified and validated indices were not as strong as anticipated, the current findings do not refute the validity of the three modified animal-based indices as on-farm measures of body condition, lameness and injury in group-housed pigs. Moderate to substantial levels of agreement demonstrated the intra- and inter-observer reliability of the three modified indices as on-farm measures of body condition score, lameness and injury in group housed pigs. Thus, the modified welfare indices had acceptable reliability in terms of intra- and inter observer agreement.

The key objectives of the project were met; examining the external validity of the modified animal-based welfare indices (body condition score, injury score and lameness score) identified by APL project 2012/1025 as on-farm measures of pig welfare in the Australian Pork Industry, and investigating the inter- and intra-observer reliability of the modified animal-based welfare indices included in the proposed on-farm pig welfare benchmarking protocol. Further research is now recommended in order to develop and trial a standardised training package to train producers in the effective on-farm application of the proposed pig welfare benchmarking protocol, at which time the additional validity testing of the three modified animal-based welfare indices could also be conducted. Confirming the validity of the three modified animal;-based welfare indices and developing a comprehensive training package are of particular importance if the protocol is to be used effectively by Australian producers to monitor and benchmark pig welfare over time, to demonstrate improvement in animal welfare outcomes over time, to identify areas of improvement for pig welfare, and to compare pig welfare across units in multisite enterprises.

Table of Contents

1. Acknowledgements	2
2. Executive Summary	3
3. Background to Research	8
1.1 Objectives of the Research Project	9
1.2 Introductory Technical Information	9
4. Research Methodology	10
5. Discussion of Results	15
6. Implications and Recommendations	23
7. Intellectual Property	25
8. Technical Summary	25
9. References	26

List of Tables

Table 1. Spearman rho correlation coefficients and Kappa statistics between validated animal-based indices and the modified animal based indices within observer, for group-housed sows	16
Table 2. Spearman rho correlation coefficients and Kappa statistics between validated animal-based indices (V) and modified animal-based indices (M) within observer, for group-housed grower pigs	17
Table 3. Intra-observer reliability between day 1 (D1) and day 2 (D2) within observer per modified animal-based welfare indice in group-housed sows	19
Table 4. Intra-observer reliability between day 1 (D1) and day 2 (D2) within observer per modified animal-based welfare indice in group-housed grower pigs	20
Table 5. Inter-observer reliability between observers on day 2 (D2) per modified animal-based welfare indice in group-housed sows	21
Table 6. Inter-observer reliability between observers on day 2 (D2) per modified animal-based welfare indice in group-housed grower pigs	21
Table 7. Inter-observer reliability between paired observers on day 1 (D1) for both validated (V) and modified (M) animal-based welfare indice in group-housed sows	22
Table 8. Inter-observer reliability between paired observers on day 1 (D1) for both validated (V) and modified (M) animal-based welfare indice in group-housed grower pigs	22

List of Figures

Figure 1. Guide for the visual and tactile body condition assessment in sows	11
Figure 2. The three body regions of the pig requiring injury scoring	12
Figure 3. The five body regions of the pig requiring injury scoring- note the front and back legs are one region.	12
Figure 4. The frequency of BCS scores (Y-axis) assigned by each observer (X-axis) on day 2 in group-housed sows (n=119)	17
Figure 5. The frequency of LS scores (Y-axis) assigned by each observer (X-axis) on day 2 in group-housed sows (n=119)	17
Figure 6. The frequency of BCS scores (Y-axis) assigned by each observer (X-axis) on day 2 in group-housed grower pigs (n=120)	18
Figure 7. The frequency of LS scores (Y-axis) assigned by each observer (X-axis) on day 2 in group-housed grower pigs (n=120)	18

3. Background to Research

The Australian Pork Industry Quality Assurance Program (APIQ) allows producers to be annually audited (PigCare audit) against a set of standards and performance indicators. This audit does not however provide producers and managers with the opportunity to monitor the welfare status of their pigs over time and/or benchmark the welfare status of animals between farms. Given that the Australian Pork Industry adopts a pro-active approach to animal welfare, a protocol which provided a simple, validated set of pig welfare indicators would both enable the industry to make evident its commitment to pig welfare and demonstrate incremental welfare improvements over time.

The preceding APL project 2012/1025, 'The development of practical measures to benchmark pig welfare in the Australian Pork Industry', identified animal-based welfare indices capable of forming a pig welfare benchmarking protocol to be used by producers to measure and monitor pig welfare on farm. Whilst the project examined the practicality of the identified animal-based welfare indices for on-farm pig welfare assessment (feasibility), the inter- and intra-observer reliability of the welfare parameters were not examined. Furthermore, in order for the benchmarking protocol to be applicable to all production systems within the Australian Pork Industry, the methodology for a number of the identified welfare indices required modification for on-farm assessment. Consequently, the identified animal-based welfare indices; body condition score, injury score and lameness score, with modified methodologies recommended for on-farm assessment require validation as on-farm measures of pig welfare.

The current project was conducted to examine the external validity and intra- and inter-observer reliability of the modified animal-based welfare indices and their modified methodologies recommended for on-farm assessment, in order to validate the indices as on-farm measures of pig welfare. If validated, the modified welfare indices (body condition score, injury score and lameness score) will be able to be included in the practical on-site pig welfare benchmarking tool proposed by the earlier APL project 2012/1025. This protocol, to be applied by farmers for self-auditing purposes, will also provide producers with the opportunity to benchmark their welfare status against others within the industry. Furthermore, such a tool may also be used by both producers and the industry to monitor pig welfare over time, to demonstrate improvement in animal welfare outcomes over time, to identify areas of improvement for pig welfare, and to compare pig welfare across units in multisite enterprises.

The current project addresses Strategy 1 under Core Objective 4 - Address Changing Expectations & Standards for Food Production and will assist Industry in further demonstrating to the community its commitment to enhancing pig welfare and in turn, it will also assist in building the confidence of consumers, retailers, government and the community that the Australian pork Industry is at the forefront of International developments in animal welfare. The project relates directly to APL Core Objective 4, (leadership, preparedness, stewardship) through Strategy 1, (address changing expectations & standards for food production) by assuring consumers, government, markets and the Australian public that the industry's ethical standards are of a very high level. It also links with APL Core Objective 5, (industry cohesion & responsiveness) by feeding into Strategy 3, (enhance the reputation & effectiveness of APL) and promoting APL as an organisation that promotes sound welfare practices. This project directly addresses the APL R&D priority to underpin welfare requirements by contributing to the development of a tool which will allow the industry to

demonstrate that the production systems it uses provide appropriate level of animal care and proper welfare above and beyond what is required under the COP.

1.1 Objectives of the Research Project

The current project sought to:

1. To examine the external validity of the modified animal-based welfare indices (body condition score, injury score, and lameness score), identified during APL Project 2012/1025, as on-farm measures of pig welfare in the Australian Pork Industry.
2. To investigate the inter- and intra-observer reliability of the modified animal-based welfare indices included in the on-site pig welfare benchmarking protocol.

1.2 Introductory Technical Information

Animal welfare, defined as an animal's 'state as regards its attempts to cope with its environment' (Broom, 1986), is a complex construct, combining both subjective and objective aspects of the animal's quality of life (Smulders et al., 2006). Despite an obvious need, a system of welfare evaluation which is accepted, comprehensive and fully-validated has yet to be developed. Animal welfare is a multidimensional concept; therefore any assessment must include resource/environment-based (input), management-based (input) and animal-based (output) measures (Fraser, 1995; Broom, 1998; Barnett and Hemsworth, 2003). Research in the field of animal welfare has become more focused and genuinely applied, with recent work focussing on the development of methods for assessing welfare in situ.

The on-farm animal welfare assessment/monitoring scheme has become an important tool for the effective management of intensively farmed animals within the different livestock industries. Furthermore, with the welfare of domestic animals high on both the political and societal agendas, an increase in pressure has shifted the scientific focus to the development of a scientifically based on-farm welfare assessment tool, able to be implemented across a wide range of domestic animal species (Winckler et al., 2003). An effective on-site animal welfare assessment and monitoring scheme needs to include both input (environment, resource and management-based) and output (animal-based) parameters which, through past research, have been shown to be valid (how relevant the measure is with regard to representing the welfare of the animal), reliable (consistency of the result with respect to sensitivity of the measure and inter- and intra-observer reliability) and feasible (the ease with which the measurement can be made with regard to time, effort and method). In recent years, the assessment and monitoring of animal welfare has shifted from the conventional approach of evaluating the environment and resources required to ensure good welfare, and instead focused on the application of animal-based measures of welfare (Knierim and Winckler, 2009).

Animal-based parameters provide a direct measurement of the animal's welfare, and while environmental parameters will offer information regarding potential or current welfare risks, they fail to directly register the state of the animal (Johnsen et al., 2001; Winckler et al., 2003). The benefit gained from directly measuring an animal's state may however be negated by the validity, reliability and feasibility limitations often associated with the measurement of animal-based parameters, which are generally greater than those associated with environmental/resource parameters (Knierim and Winckler, 2009). Although the assessment of animal welfare at a farm level remains an on-going challenge for animal welfare scientists, the scientific literature demonstrates the opportunity to develop a practical and effective on-site welfare assessment tool, using validated, repeatable and

feasible animal-based welfare indices, capable of benchmarking pig welfare in the Australian Pork Industry. Whilst they have been extensively employed within the literature and more recently within on-site assessment schemes, the validity of animal-based welfare indices with regard to pigs has yet to receive meaningful examination.

A number of the animal-based welfare indices identified by the preceding APL project 2012/1025, for inclusion in an on-farm pig welfare benchmarking protocol to be used by producers to measure and monitor pig welfare on farm, required methodology modification for on-farm assessment. Consequently, the identified animal-based welfare indices; body condition score, injury score and lameness score, with modified methodologies recommended for on-farm assessment required validation as on-farm measures of pig welfare.

Therefore, the current project aimed to;

1. Examine the external validity of the modified animal-based welfare indices, body condition score, injury score and lameness score, which were identified by APL project 2012/1025 as on-farm measures of pig welfare in the Australian Pork Industry, and
2. Investigate the inter- and intra-observer reliability of the modified animal-based welfare indices included in the proposed on-farm pig welfare benchmarking protocol.

4. Research Methodology

The external validity and inter- and intra-observer reliability of the modified animal-based welfare indices, body condition score, injury score and lameness score, were investigated in group-housed sows and grower pigs at a large commercial piggery in Regional Victoria, over a 6-week period in June and July, 2014. Sampling was conducted by four trained observers (Obs1, Obs2, Obs3 and Obs4); each observer assessed 20 animals for BCS, IS, and LS on day 1 of the sampling period using both the validated (V) and the modified (M) methodologies, and 40 animals using the modified (M) methodologies on day 2 (consecutive days). Two hundred and forty animals were sampled; 120 sows in weeks 1-3 and 120 grower pigs in weeks 4-6 of the project. Forty animals were sampled per day; sampling was conducted over six 2-day periods (3 two-day periods for sows and another 3 two-day periods for grower pigs). On day 0 of each week a stockperson would select and mark up 40 animals; animals were selected to cover a range of conditions, with regard to body condition, gait soundness, and injuries in an attempt to obtain a sample with sufficient variation.

The external validation of the modified animal-based welfare indices: animals will be assessed using the “validated” and “modified” tests for Body Condition Score, Injury Score and Lameness Score

To examine the validity of the modified animal-based welfare indices as on-farm measures of pig welfare, comparisons were made between scores obtained from the modified methodologies and scores gained from those methodologies which had previously been validated under experimental conditions, in order to determine the level of agreement between the two measures. The modified animal-based welfare indices which were examined in group housed sows and grower pigs were:

Body Condition Score (BCS): the modified 3-point scoring scale recommended for on-farm assessment of BCS was compared with Patience & Thacker’s (1989) 5-point scoring scale.

Modified BCS: a visual BCS assessment of the animal from in front, behind and alongside, using the simplified 3-point scoring scale;

1. Thin: hips and spin are visible and animal appears very thin/thin (corresponds to Patience & Thacker’s (1989) BCS 1 and 2)

2. Normal: hips and spine lack visual prominence (corresponds to Patience & Thacker's (1989) BCS 3 and 4)
3. Overfat: hips and spine are heavily covered and animal appears visually obese (corresponds to Patience & Thacker's (1989) BCS 5)

Validated BCS (Patience & Thacker, 1989): a visual and tactile assessment of the animal from in front, behind and alongside using the 5-point scoring scale (Figure 1);

1. Emaciated: hips and spine are visually prominent; pig appears very thin
2. Thin: hips and spine visible and are easily felt without pressure applied through the palm
3. Ideal: hips and spine lack visual prominence and are only felt when pressure is applied through the palm
4. Fat: hips and spine are not visible and cannot be felt when firm pressure is applied through the palm
5. Overfat: hips and spine are heavily covered and cannot be felt even when pressure is applied with a single finger; sow appears visually obese

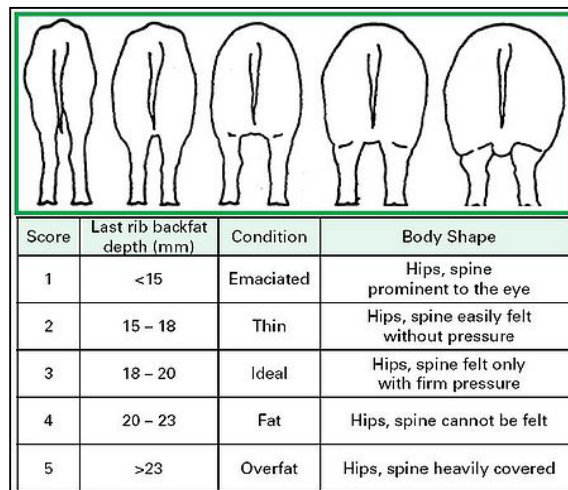


Figure 1. Guide for the visual and tactile body condition assessment in sows

Injury Score (IS): comparisons will be made between the modified methodology recommended for on-farm assessment of IS in pigs and the common 4-point scoring scale derived from the well-established De Koning scale (De Koning, 1984).

Modified IS: a visual assessment of one side of the animal is conducted using a simplified count of all lesions over three sections of the animal's body (Figure 2); head/shoulders (including ears), middle and rear. A count of all injuries in each section of the animal will be made and the injury count will be classified as;

Injury count: a lesion count using a 4-point scale,

0. no scratches/cuts/other lesions,
1. < 5 scratches/cuts/other lesions,
2. 5-9 scratches/cuts/other lesions,
3. >10 scratches/cuts/other lesions, and

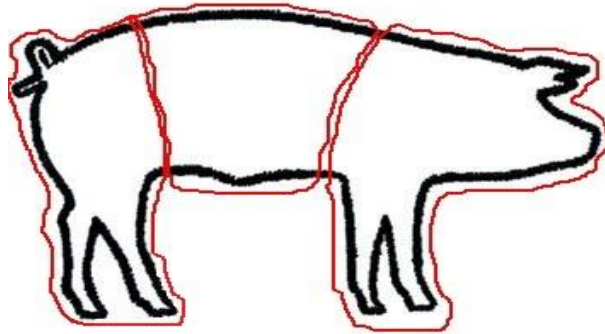


Figure 2. The three body regions of the pig requiring injury scoring

Validated IS (De Koning, 1984): a visual assessment of one side of the animal for fresh injury and wounds, whereby five specified regions of the animal's body (Figure 3) are assigned an injury score using a 4-point scoring scale derived from the well-established De Koning scale;

0. no lesions,

1. 1-5 lesions,

2. 6-10 lesions, and

3. >10 lesions

3. In order to standardise assessment;

- If there is less than 0.5cm between two lesions they will be measured as one lesion (i.e. a LS of 1)
- A group of small lesions with less than 0.5cm between them are scored according to the size of the area; diameter <2 cm equals one lesion (LS of 1), diameter between 2 and 5 cm equals six lesions (LS of 2), diameter >5 cm equals ten lesions (LS of 3)
- A bleeding lesion between 2 and 5cm, or a healed lesion of more than 5cm will be considered as six lesions (i.e. a LS of 2)
- A deep and open lesion of more than 5cm will be considered as ten lesions (LS of 3)

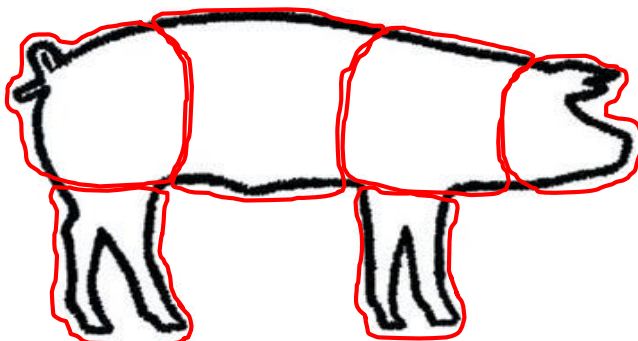


Figure 3. The five body regions of the pig requiring injury scoring- note the front and back legs are one region.

Injuries were deemed to be fresh according to their colour (pink/red hue around injury) and the presence of a 'pickable' scab.

Lameness Score (LS): the modified 3-point scale recommended for on-farm lameness assessment in pigs will be compared with Karlen et al. (2007) validated 4-point assessment scale.

Modified LS: a visual assessment of the pig from behind and alongside, whilst walking for at least 5 clear strides. If the pig is lying down, the observer forces the animal to rise and allows it to take at least 5 clear strides until balanced before then performing the assessment. The assessment of LS is performed using a 3-point scale (modified from Karlen et al., 2007);

0. Sound: a regular gait; the animal's ability to stand and move is not affected and their limb movements are symmetrical (corresponds to Karlen et al (2007) LS of 0)
1. Lamé: a visible degree of difficulty in walking or a visible reluctance to bear weight on the affected limb, but still using all four legs; swagger of caudal body while walking, shortened stride (corresponds to Karlen et al. (2007) LS of 1 and 2)
2. Non-weight bearing: no weight bearing on affected limb; total recumbency (corresponds to Karlen et al. (2007) LS of 3)

Validated LS (Karlen et al., 2007): a visual assessment of the animal from in front, behind and alongside, whilst walking, trotting and standing. The assessment is conducted by two experimenters (one to move the animal and one to perform the assessment) and all animals are allowed to walk freely for at least 30m before the assessment is performed. Each animal is forced to walk and trot at least 50m along a concrete walkway (test is stopped if animal is found to be severely lame), and lameness is assessed using a 4-point scale;

0. Sound: the animal's ability to stand and move is not affected and their limb movements are symmetrical
1. Compromised movement: the animal is not considered lame in that it appears their ability to stand and move is unaffected and all their legs are able to bear weight similarly, but their movement is compromised
2. Moderately lame: the animal's ability to stand is obviously reduced, movement is diminished and difficult, and they refuse to bear weight on one or more limbs
3. Severely lame: the animal's ability to stand and move is very restricted, they are unable to bear weight on one or more limbs, and swollen joints, stiffness and vocalisations are often observed when they are moved

The inter- and intra-observer reliability of the animal-based welfare indices:

In order to be considered a valid measure of on-farm pig welfare, the scoring of the animal-based welfare indices needs to be consistent (reliable) when being applied repeatedly by the same assessor (intra-observer) or independently by different assessors (inter-observer). The inter- and intra-observer reliability of the animal-based welfare indices (BCS, IS, LS) was investigated using a test-retest assessment (Whitham and Wielebnowski, 2009), which examined the similarity between measures collected on an animal at two different time points (intra-observer) or the agreement between measures taken on an animal by multiple assessors (inter-observer).

Methodology: Investigating the validity and reliability of the modified animal-based welfare indices in group-housed pigs (3 two-day replicates for sows and 3 two-day replicates for growers)

Day 0

Stock person selected and marked up 40 animals ranging in condition, with regard to body condition, gait soundness, and injuries.

Day 1

40 group-housed animals were assessed for injuries, lameness and body condition, using both the validated (V) and the modified (M) methodologies. Assessment of the validated indices was performed in individual animals in an empty group pen and a secured laneway, whilst the modified indices were performed on animals in their group pen.

- Observers 1 and 2 assessed animals 1-20 using both the validated and the modified animal-based welfare indices, alternating the order of assessment; animals 1-10 were assessed using the V methodology first followed by the M methodology, whilst animals 11-20 were assessed using the M methodology first followed by the V methodology.
- Observers 3 and 4 assessed animals 21-40 using both the validated and the modified animal-based welfare indices; alternating the order of assessment, animals 21-30 were assessed using the V methodology first followed by the M methodology, whilst animals 31-40 were assessed using the M methodology first followed by the V methodology.
- Observers and the pairing of observers were kept consistent across the study, however the animal numbers assessed by each pair on day 1 were alternated, i.e. on week 1 Observers 1 and 2 assessed animals 1-20 and Observers 3 and 4 assessed animals 21-40, whilst on week 2 Observers 1 and 2 assessed animals 21-40 and Observers 3 and 4 assessed animals 1-20, and so forth for the remainder of the study.

Day 2

40 group housed animals will be assessed for a second time for injuries, lameness and body condition, using the modified methodologies. Animals were assessed in their group pen.

- Observer 1, 2, 3 and 4 assessed animals 1-40 using the modified animal-based welfare indices.

NOTE: The reliability of the validated human withdrawal test (HWT) described by Hemsworth et al. (1981) was not examined as originally proposed. The HWT (Hemsworth et al., 1981), which measures the withdrawal response of lactating sows to an observer approaching in a standard manner, is conducted by two observers and as such is not suitable for inclusion in an on-farm pig welfare benchmarking tool to be applied by a single observer. To date, tests capable of measuring a pig's fear of humans (and subsequently the human-animal relationship) in lactating and group-housed animals have yet to be fully developed and validated for on-farm use.

The validity of the modified animal-based welfare indices as on-farm measures of pig welfare were investigated using Spearman's rho (ρ) correlations to examine the strength of the relationship between the modified and the validated animal-based indices, and Kappa statistics (κ) were calculated to examine the level of agreement between the validated and the modified animal-based indices. Intra-observer assessment was investigated using Spearman's rho (ρ) correlations to examine the strength of relationship between an observer's assessment at two different time points, and their

repeated assessment of the same animal on day 2, for each of the three modified animal-based welfare indices. In addition, Kappa statistics (κ) were calculated to examine the level of agreement between the observer's indice scores on day 1 and day 2. Inter-observer assessment was investigated using Spearman's rho (ρ) correlations to examine the relationship between the modified indice scores taken by multiple observers on the same animals at the same time point, for each of the three modified animal-based welfare indices. The level of agreement between the multiple observers modified indices scores on the same animal on day 2 was then examined using Kappa statistics (κ).

Categories of ρ and κ value are usually assigned a description in order to better understand the strength of a relationship or the level of agreement between measures. A commonly used scale to interpret Spearman rho correlations is small/low $\rho = 0.10$ to 0.29 , medium/moderate $\rho = 0.30$ to 0.49 , and large/high $\rho = 0.50$ to 1.0 (Cohen, 1988). In addition, Martin and Bateson's (2007) general guideline that correlations of 0.7 or greater are acceptable because it indicates that approximately half of the variation in the second set of scores is explained by the first set. For Kappa statistics, Landis and Koch (1977) suggest that < 0.00 κ : poor agreement, $0.00-0.20$ κ : slight agreement, $0.21-0.40$ κ : fair agreement, $0.41-0.60$ κ : moderate agreement, $0.61-0.80$ κ = substantial agreement, $0.81-1.00$ κ = almost perfect agreement. Cohen's (1988) ρ scale and Landis and Koch's (1977) κ scales were used to interpret the results of this project.

5. Discussion of Results

The external validation of the modified animal-based welfare indices: animals will be assessed using the “validated” and “modified” tests for Body Condition Score, Injury Score and Lameness Score

The validity of the modified animal-based welfare indices as on-farm measures of pig welfare was investigated by comparing the assessment scores obtained from the modified methodologies and those from the validated methodologies, in order to determine the level of agreement between the two measures. Spearman's rho (ρ) correlations were used to examine the strength of the relationship between the modified and the validated animal-based indices. A total IS was calculated for each animal by summing the individual IS scores from each of the body sections (the five sections for the validated IS and the three sections for the modified IS). Following a re-coding (adjustment) of the validated indice data, in order to allow for the comparison of the indices with different scoring scales, Kappa statistics (κ) were calculated to examine the level of agreement between the validated and the modified animal-based indice scores within observers. Spearman's rho (ρ) correlations and Kappa statistics (κ) for sows are given in Table 1 and for grower pigs in Table 2.

In group-housed sows (Table 1), correlations between the validated and the modified animal-based welfare indices were significant (with the exception of observer 2 for BCS); however, they tended to be low to moderate in strength. These findings were supported by the Kappa statistics which demonstrated slight to fair agreement between the validated and modified animal-based welfare indices for body condition score and lameness score, and poor agreement between the two indices with regard to injury score. For group-housed grower pigs (Table 2) the correlations between the validated and the modified indices were significant, and the strength of the correlations ranged from moderate to high. With regard to agreement between the validated and the modified indices, the Kappa statistics varied from fair through to a substantial level of agreement. The poor to fair agreement between the validated and the modified IS may be explained by the fact that two different

degrees of injury are being assessed, i.e. the validated indice is only measuring the degree of fresh injuries, whilst the modified indice is measuring the degree of total injury (all injuries).

In order to demonstrate that the modified animal-based welfare indices are valid on-farm measures of body condition, lameness and injury in group-housed pigs, high correlations (greater than 0.50 ρ) and at least moderate agreement (greater than 0.60 κ) between the validated and the modified measures were sought. The modified animal-based indices were derived from validated methods of body condition, lameness and injury assessment, based on the scientific literature and practical plausibility. This was achieved by simplifying the scoring scales and reducing the subjectivity of the assessment in order to improve their on-farm practicality. Given that the validated animal-based welfare indices underwent only minor modification, it was anticipated that the modified indices would correlate more highly and have better levels of agreement with the validated measures than the results suggest. One explanation for this finding is the homogeneity of the data. The lack of variation in the condition of the animals sampled is evident in Figures 4 and 5, which show the frequency of indice scores assigned by each observer on day 2, for BCS and LS respectively. Similarly, the findings for grower pigs are shown in Figures 6 and 7. The assessment of animal-based welfare indices is somewhat subjective, and as such a small degree of inconsistency in observer perception and interpretation is normal (Meagher, 2009). The lack of variation in the current sample means that these type of minor inconsistencies between the validated and the modified indices are enough to substantially reduce the statistical level of agreement between the measures. Whilst the statistical relationship between the modified and validated indices is not as strong as would be liked, the intra- and inter- reliability between observers was reasonable despite the lack of variation. Given these findings it would be expected that further validity testing, performed in populations with greater variation (i.e. at multiple sites which display a greater range of welfare outcomes, including the extremes), may well confirm the validity of the three modified animal-based indices. Further validity testing should be conducted before the three modified indices are made available to Australian producers. A comprehensive training package to train producers in the effective on-farm application of the pig welfare benchmarking protocol will need to be developed and trialed, at which time additional validity testing of the three animal-based welfare indices could also be performed.

Table 1. Spearman rho correlation coefficients and Kappa statistics between validated animal-based indices and the modified animal based indices within observer, for group-housed sows

Animal welfare indice	Observer			
	1 (V-M)	2 (V-M)	3 (V-M)	4 (V-M)
Spearman's rho				
Body condition score	0.263* (n=60)	0.227 (n=60)	0.357** (n=60)	0.356** (n=60)
Lameness score	0.268* (n=60)	0.374** (n=60)	0.448** (n=60)	0.356** (n=60)
Injury score	0.511** (n=60)	0.33** (n=60)	0.296* (n=60)	0.298* (n=60)
Kappa statistic				
Body condition score	0.487** (n=60)	0.353** (n=60)	0.413** (n=60)	0.308** (n=60)
Lameness score	0.206** (n=60)	0.302** (n=60)	0.173** (n=60)	0.205** (n=60)
Injury score	0.063 (n=60)	0.05 (n=60)	0.058 (n=60)	0.022 (n=60)

* $p < 0.05$, ** $p < 0.001$

Table 2. Spearman rho correlation coefficients and Kappa statistics between validated animal-based indices (V) and modified animal-based indices (M) within observer, for group-housed grower pigs

Animal welfare indice	Observer			
	1 (V-M)	2 (V-M)	3 (V-M)	4 (V-M)
Spearman's rho				
Body condition score	0.307** (n=60)	0.689** (n=60)	0.452** (n=60)	0.382** (n=60)
Lameness score	0.850** (n=60)	0.944** (n=60)	0.701** (n=60)	0.566 (n=60)
Injury score	0.568** (n=60)	0.486** (n=60)	0.289* (n=60)	0.289* (n=60)
Kappa statistic				
Body condition score	0.412** (n=60)	0.576** (n=60)	0.227* (n=60)	0.314** (n=60)
Lameness score	-0.007 (n=60)	0.942** (n=60)	0.689** (n=60)	0.526* (n=60)
Injury score	0.282** (n=60)	0.279** (n=60)	0.128 (n=60)	0.041 (n=60)

*p < 0.05, ** p < 0.001

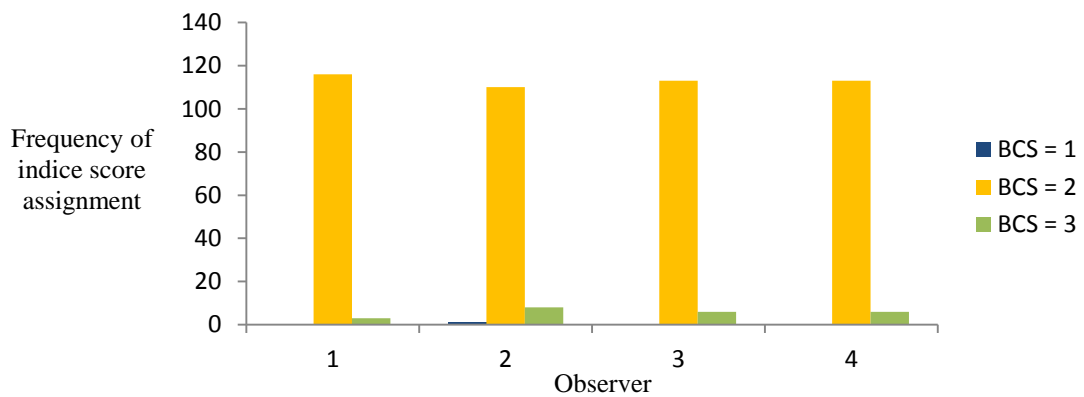


Figure 4. The frequency of BCS scores (Y-axis) assigned by each observer (X-axis) on day 2 in group-housed sows (n=119)

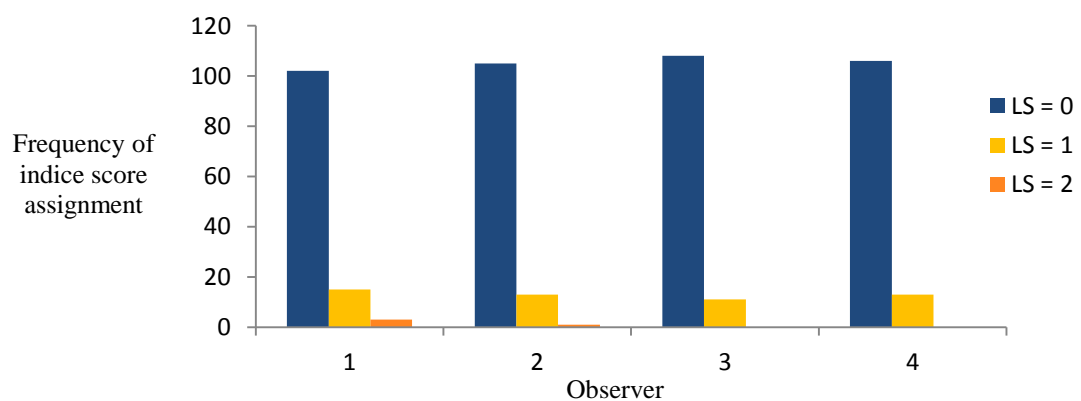


Figure 5. The frequency of LS scores (Y-axis) assigned by each observer (X-axis) on day 2 in group-housed sows (n=119)

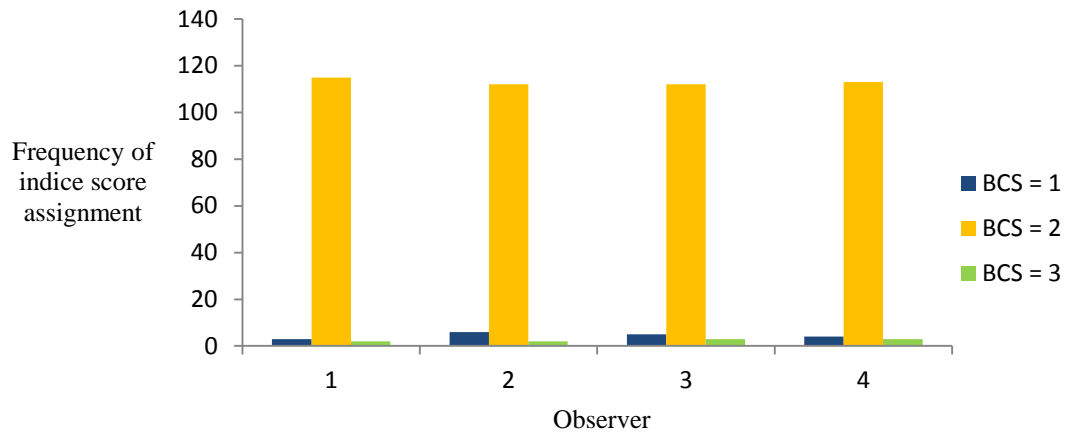


Figure 6. The frequency of BCS scores (Y-axis) assigned by each observer (X-axis) on day 2 in group-housed grower pigs (n=120)

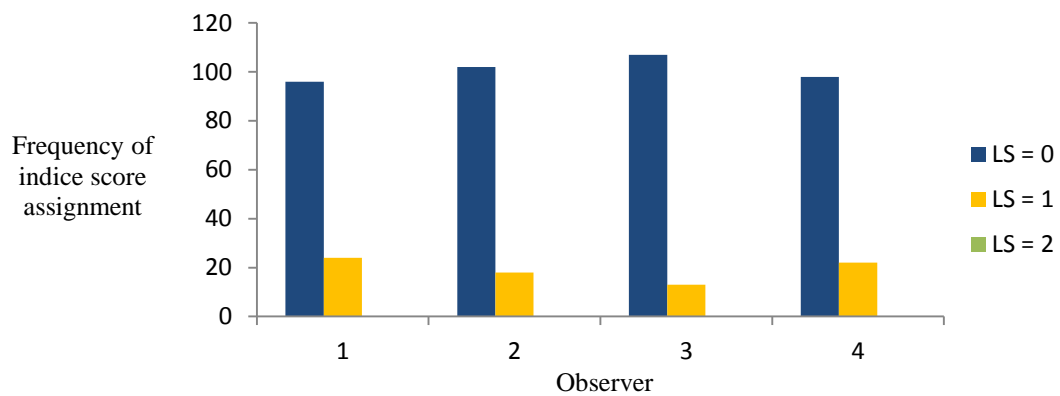


Figure 7. The frequency of LS scores (Y-axis) assigned by each observer (X-axis) on day 2 in group-housed grower pigs (n=120)

The inter- and intra-observer reliability of the animal-based welfare indices:

The inter- and intra-observer reliability of the modified animal-based welfare indices (BCS, IS, LS) was investigated using a test-retest assessment (Whitham and Wielebnowski, 2009), which examined the similarity between measures collected on an animal at two different time points (intra-observer) and the agreement between measures taken on an animal by multiple assessors (inter-observer). The test-retest assessment for intra-observer assessment was investigated using Spearman's rho (ρ) correlations to examine the strength of relationship between an observer's assessment at two different time points, i.e. the relationship between an observer's assessment of an animal on day 1 and their repeated assessment of the same animal on day 2, for each of the three modified animal-based welfare indices. In addition, Kappa statistics (κ) were calculated to examine the level of agreement between the observer's indice scores on day 1 and day 2. The Spearman's rho (ρ) correlations and Kappa statistics (κ) for sows are given in Table 3 and for grower pigs in Table 4. The test-retest assessment for inter-observer assessment was investigated using Spearman's rho (ρ) correlations to examine the relationship between the modified indice scores taken by multiple observers on the same animals at the same time point, i.e. the relationship between observers' assessments of the same animal on day 2, for each of the three modified animal-based welfare

indices. The level of agreement between the multiple observers modified indices scores on the same animal on day 2 was then examined using Kappa statistics (κ). The Spearman's rho (ρ) correlations and Kappa statistics (κ) for sows are given in Table 5 and for grower pigs in Table 6.

Intra-observer correlations were significant for all three of the modified animal-based welfare indices in group-housed sows (Table 3). The strength of the intra-observer correlations ranged from moderate to high; BCS on day 1 tended to correlate highly with BCS on day 2, day 1 LS correlated moderately with LS on day 2, and IS on day 1 correlated highly with IS on day 2. The Kappa statistics were all significant, however the intra-observer agreement between measures on day 1 and day 2 varied from fair agreement to an almost perfect level of agreement. The level of intra-observer agreement was fair to moderate in both BCS and LS, and very good in IS. Whilst these findings largely suggest an acceptable degree of intra-observer reliability in BCS (moderate) and IS (substantial), the intra-observer reliability in LS in group-housed sows is only fair. This LS finding is consistent with other research, such as D'Eath (2012). The visual assessment of lameness in group-housed animals, perhaps more so than most other animal-based welfare indices, is recognised as being a subjective measure with poor reliability, which depends mainly on the training or the experience of the observer (Main et al., 2000; O'Callaghan et al., 2003). The difficulty involved with quantifying lameness in animals on-farm may be due to the difficulty associated with obtaining an adequate observation of the animal walking unhindered on a solid and consistent surface, and the difficulty observers often have distinguishing between a normal gait (LS = 0) and an irregular gait (LS = 1) (Kaler et al., 2009). A number of studies, involving a range of different animal species, have suggested that ways of improving on-farm intra-observer reliability may include further simplifying the scoring system (lame vs non-lame) and increasing the observers experience and training (Bokkers et al., 2009; Brenninkmeyer et al., 2007; D'Eath, 2012; Rutherford et al., 2009).

Intra-observer correlations were significant for all of the modified animal-based indices in group-housed grower pigs (Table 4). BCS, LS and IS on day 1 were highly correlated with BCS, LS and IS on day 2. Similarly, the Kappa statistics demonstrated substantial levels of intra-observer agreement for all of the modified animal-based welfare indices. High intra-observer correlations (greater than 0.50 ρ) and substantial levels of intra-observer agreement (greater than 0.60 κ) clearly demonstrate the on-farm intra-observer reliability in BCS, LS, and IS in group-housed growers pigs.

Table 3. Intra-observer reliability between day 1 (D1) and day 2 (D2) within observer per modified animal-based welfare indice in group-housed sows

Animal welfare indice	Observer			
	1 (D1-D2)	2 (D1-D2)	3 (D1-D2)	4 (D1-D2)
Spearman's rho				
Body condition score	0.567** (n=60)	0.484** (n=60)	0.587** (n=60)	0.577** (n=60)
Lameness score	0.407** (n=60)	0.457** (n=60)	0.662** (n=60)	0.405** (n=60)
Injury score	0.845** (n=60)	0.903** (n=60)	0.906** (n=60)	0.897** (n=60)
Kappa statistic				
Body condition score	0.487** (n=60)	0.466** (n=60)	0.492** (n=60)	0.489** (n=60)
Lameness score	0.340* (n=60)	0.446** (n=60)	0.362** (n=60)	0.268* (n=60)
Injury score	0.694** (n=60)	0.788** (n=60)	0.887** (n=60)	0.865** (n=60)

* $p < 0.05$, ** $p < 0.001$

Table 4. Intra-observer reliability between day 1 (D1) and day 2 (D2) within observer per modified animal-based welfare indice in group-housed grower pigs

Animal welfare indice	Observer			
	1 (D1-D2)	2 (D1-D2)	3 (D1-D2)	4 (D1-D2)
Spearman's rho				
Body condition score	0.77** (n=60)	0.908** (n=60)	0.894** (n=60)	0.823** (n=60)
Lameness score	0.783** (n=60)	0.877** (n=60)	0.745** (n=60)	0.796** (n=60)
Injury score	0.922** (n=60)	0.824** (n=60)	0.901** (n=60)	0.898** (n=60)
Kappa statistic				
Body condition score	0.737** (n=60)	0.901** (n=60)	0.882** (n=60)	0.793** (n=60)
Lameness score	0.778** (n=60)	0.870** (n=60)	0.739** (n=60)	0.795** (n=60)
Injury score	0.845** (n=60)	0.722** (n=60)	0.822** (n=60)	0.861** (n=60)

*p < 0.05, ** p < 0.001

Inter-observer correlations were significant for all of the modified animal-based welfare indices in group-housed sows (Table 5). The strength of these correlations ranged from moderate to high; BCS tended to correlate highly between observers, LS generally correlated moderately between observers, and IS correlated highly between observers. All Kappa statistics were significant. The level of inter-observer agreement varied from moderate to substantial. In group-housed growers, inter-observer correlations were significant for all of the modified animal-based indices (Table 6), with the exception of that between observer 1 and observer 4 in BCS. The strength of the correlations tended to be high, however there were a number of moderate inter-observer correlations largely concerning pairings involving observer 1. The Kappa statistics were for the most part significant and indicated a moderate to substantial level of agreement between observers. Interestingly, observer 1 had poor agreement with all other observers with regard to BCS. Whilst these findings suggest acceptable inter-observer reliability in the three modified animal-based welfare indices in group-housed grower pigs, there was variation in reliability amongst the observer pairings. The observed variation in results between observer pairings is likely due to the homogeneity of the data and the subjective nature of measuring animal-based welfare indices.

The intra- and inter-observer reliability of the three modified indices was commonly found to be moderate to substantial, which is consistent with acceptable levels of reliability reported in the literature. Therefore, these findings demonstrate the intra- and inter-observer reliability of the modified animal-based welfare indices.

Interestingly, the findings discussed above suggest that the validity and the reliability of the modified animal-based indices are greater in group-housed growers than in group-housed sows. A possible explanation for this finding may be observer experience. It has been widely recognized that animal-based welfare indices can be highly subjective and require observers to have both appropriate training and experience in order to be applied effectively (Knierim and Winckler, 2009; Winckler et al., 2003). Whilst the four observers had undergone basic training and had experience with animal-based welfare indices before the study commenced, their experience on-farm obviously increased as data collection occurred. The validity and reliability of the animal-based welfare indices was examined in group-housed sows in weeks 1-3, and in group-house growers in weeks 4-6 of the project. As such, it is possible that a greater observer experience when assessing grower pigs could have reduced the subjectivity of the modified indices and resulted in a greater validity and reliability

of BCS, LS, and IS in grower pigs when compared with sows. In addition, this finding may suggest that the basic training the observers undertook was insufficient to effectively employ the modified animal-based welfare indices. Thus a more comprehensive and standardised training program, which identifies the full range of welfare outcomes, may be necessary to effectively use the modified animal-based welfare indices to measure body condition, lameness and injury in group-housed pigs.

Table 5. Inter-observer reliability between observers on day 2 (D2) per modified animal-based welfare indice in group-housed sows

Animal welfare indice	Observers					
	I and 2 (n=120)	I and 3 (n=120)	I and 4 (n=120)	2 and 3 (n=120)	2 and 4 (n=120)	3 and 4 (n=120)
Spearman's rho						
Body condition score	0.567**	0.453**	0.453**	0.525**	0.669**	0.649**
Lameness score	0.462**	0.370**	0.397**	0.419*	0.535**	0.725**
Injury score	0.752**	0.792**	0.823**	0.733**	0.767**	0.853**
Kappa statistic						
Body condition score	0.481**	0.425**	0.425**	0.505**	0.646**	0.649**
Lameness score	0.415**	0.321**	0.359**	0.422**	0.543**	0.722*
Injury score	0.476**	0.634**	0.686**	0.273*	0.393**	0.769

*p < 0.05, ** p < 0.001

Table 6. Inter-observer reliability between observers on day 2 (D2) per modified animal-based welfare indice in group-housed grower pigs

Animal welfare indice	Observers					
	I and 2 (n=120)	I and 3 (n=120)	I and 4 (n=120)	2 and 3 (n=120)	2 and 4 (n=120)	3 and 4 (n=120)
Spearman's rho						
Body condition score	0.311**	0.313**	0.166	0.876**	0.803**	0.799**
Lameness score	0.607**	0.295*	0.678**	0.304**	0.706**	0.458**
Injury score	0.730**	0.786**	0.778**	0.668**	0.718**	0.811**
Kappa statistic						
Body condition score	0.011	0.012	0.020	0.868**	0.790**	0.790**
Lameness score	0.598**	0.277*	0.677**	0.299**	0.701**	0.438**
Injury score	0.434**	0.616**	0.613**	0.345**	0.436**	0.669**

*p < 0.05, ** p < 0.001

Whilst it could be interpreted that the inconclusive results of the validity testing disproves the validity of the modified animal-based welfare indices as on-farm measures of pig welfare, given that they stem from minor modifications of the validated indice, the findings are more likely due to the homogeneity of the data resulting from a lack of variation in the sample, rather than a genuine lack of validity and reliability. This explanation is supported by examining the inter-observer reliability between paired observers on day 1 for both the validated and the modified animal-based welfare

indices (Tables 7 and 8). The results demonstrate inter-observer reliability, with comparable levels of agreement, in both the validated and the modified animal-based indices in group-housed sows and grower pigs. If anything, the strength of the correlation coefficients and the level of agreement between the paired observers are greater for the modified indices than the validated ones. These results therefore support the summation that the projects findings are more likely a result of the homogeneity of the data rather than an actual lack of validity and reliability in the three modified animal-based welfare indices. That said, current findings support further investigation in order to confirm the validity of the modified animal-based welfare indices before they are made available to Australian pig producers. Furthermore, the variability of the results reiterates the subjective nature of animal-based welfare indices and emphasises the need to develop a comprehensive standardised training package for the on-farm pig welfare benchmarking protocol.

Table 7. Inter-observer reliability between paired observers on day 1 (D1) for both validated (V) and modified (M) animal-based welfare indice in group-housed sows

Animal welfare indice	Observer			
	I and 2 (V)	I and 2 (M)	3 and 4 (V)	3 and 4 (M)
Spearman's rho				
Body condition score	0.49** (n=60)	0.487** (n=60)	0.504** (n=60)	0.673** (n=60)
Lameness score	0.562** (n=60)	0.586** (n=60)	0.484** (n=60)	0.785** (n=60)
Injury score	0.791** (n=60)	0.831** (n=60)	0.810** (n=60)	0.622** (n=60)
Kappa statistic				
Body condition score	0.279* (n=60)	0.384** (n=60)	0.376** (n=60)	0.654** (n=60)
Lameness score	0.521** (n=60)	0.562** (n=60)	0.399** (n=60)	0.748** (n=60)
Injury score	0.540** (n=60)	0.596** (n=60)	0.616** (n=60)	0.723** (n=60)

*p < 0.05, ** p < 0.001

Table 8. Inter-observer reliability between paired observers on day 1 (D1) for both validated (V) and modified (M) animal-based welfare indice in group-housed grower pigs

Animal welfare indice	Observer			
	I and 2 (V)	I and 2 (M)	3 and 4 (V)	3 and 4 (M)
Spearman's rho				
Body condition score	0.687** (n=60)	0.701** (n=60)	0.966** (n=60)	0.638** (n=60)
Lameness score	0.901** (n=60)	0.880** (n=60)	0.721** (n=60)	0.856** (n=60)
Injury score	0.697** (n=60)	0.812** (n=60)	0.797** (n=60)	0.800** (n=60)
Kappa statistic				
Body condition score	0.654** (n=60)	0.648** (n=60)	0.920** (n=60)	0.554** (n=60)
Lameness score	0.896** (n=60)	0.880** (n=60)	0.707** (n=60)	0.856** (n=60)
Injury score	0.510** (n=60)	0.548** (n=60)	0.694** (n=60)	0.707** (n=60)

*p < 0.05, ** p < 0.001

The key objectives of the project were met; examining the external validity of the modified animal-based welfare indices (body condition score, injury score and lameness score) identified by APL project 2012/1025 as on-farm measures of pig welfare in the Australian Pork Industry, and

investigating the inter- and intra-observer reliability of the modified animal-based welfare indices included in the proposed on-farm pig welfare benchmarking protocol. Validity testing largely indicated a moderate relationship with slight to fair levels of agreement between the validated and the modified animal-based welfare indices. This finding is likely due to the homogeneity of the data, and as such further testing could be performed in populations with greater variation in order to confirm the validity of the modified animal-based indices. Whilst the correlations and the level of agreement between the modified and validated indices were not as strong as anticipated, the current findings do not refute the validity of the three modified animal-based indices as on-farm measures of body condition, lameness and injury in group-housed pigs. The intra- and inter-observer reliability of the three modified indices was largely found to be moderate to substantial, which is consistent with acceptable levels of reliability reported in the literature. Therefore, these findings demonstrate the intra- and inter-observer reliability of the modified animal-based welfare indices. Further research is now recommended in order to develop and trial a standardised training package to train producers in the effective on-farm application of the proposed pig welfare benchmarking protocol, at which time the additional validity testing of the three modified animal-based welfare indices could also be conducted. Confirming the validity of the three modified animal-based welfare indices and developing a comprehensive training package are of particular importance if the protocol is to be used effectively by Australian producers to monitor and benchmark pig welfare over time, to demonstrate improvement in animal welfare outcomes over time, to identify areas of improvement for pig welfare, and to compare pig welfare across units in multisite enterprises.

The results of this project will be reported to the Australian pig industry via the APL Final Report, targeted industry meetings, targeted industry publications, targeted meetings with industry consultants and extension personnel and scientific publications.

6. Implications and Recommendations

The current project examined the validity and the inter- and intra-observer reliability of the modified animal-based welfare indices, body condition score, injury score and lameness score, identified by APL project 2012/1025 as on-farm measures of welfare in group-housed pigs, which could potentially be included in a proposed on-farm pig welfare benchmarking protocol, to be applied by Australian producers for self-auditing purposes. Whilst the current findings do not refute the validity of the three modified animal-based indices as on-farm measures of body condition, lameness and injury in group-housed pigs, they do suggest that further validity testing is warranted in order to confirm the validity of the indices before they are made available to the industry. Lameness, injury and inappropriate body condition are all important issues which have the potential to impact significantly on pig health, welfare and production. Consequently, in order to be effective any on-farm pig welfare benchmarking protocol must include valid measures of body condition, lameness and injury (BCS, LS and IS). The modified animal-based welfare indices investigated in this study were derived from validated methods of body condition, lameness and injury assessment, according to the scientific literature and practical plausibility. Modifications involved simplifying the scoring scales and reducing the subjectivity of the assessment in order to improve their on-farm practicality. Given that the validated animal-based welfare indices underwent only minor modification, we still maintain that the modified indices (BCS, LS, and IS) are valid and reliable measures of body condition, lameness and injury in group-housed pigs. The projects findings may be explained by the homogeneity of the data resulting from a lack of variation in the condition of the animals sampled, rather than a genuine lack of validity and reliability of the three modified animal-based welfare indices. Attempts to obtain a varied sample were made by selecting animals which ranged in condition; however this proved

unsuccessful because the sampled population lacked variation and extremes with regard to animal condition. Consequently, any further investigation of the validity of the three modified animal-based welfare indices needs to be conducted in populations with greater variation, i.e. multiple sights which display a greater range of welfare outcomes (including the extremes). Furthermore, a comprehensive and standardised training package to train producers in the effective on-farm application of the pig welfare benchmarking protocol will need to be developed and trialed, at which time the additional validity testing of the three animal-based welfare indices could also be performed. The authors believe that a sample with variation in welfare outcomes will enable the validity and reliability of the modified animal-based welfare indices as on-farm measures of body condition, lameness and injury in group-housed sows, to be clearly demonstrated. A fully developed on-farm pig welfare benchmarking protocol, including valid BCS, LS and IS, will enable producers and the industry to monitor pig welfare over time, to demonstrate improvement in animal welfare outcomes over time, to identify areas of improvement for pig welfare, and to compare pig welfare across units in multisite enterprises.

The current project had originally intended to examine the reliability of the validated human withdrawal test (HWT) described by Hemsworth et al. (1981). However, given that the HWT requires two observers it was deemed unsuitable for inclusion in an on-farm pig welfare benchmarking tool to be applied by a single observer. To date, tests capable of measuring a pig's fear of humans (and subsequently the human-animal relationship) in lactating and group-housed animals have yet to be fully developed and validated for on-farm use. A measure of the human-animal relationship (HAR) is recommended for inclusion in the proposed on-farm pig welfare benchmarking protocol. The scientific literature indicates that the HAR, measured by a fear of humans, may be an important indicator of animal welfare; however it remains a challenge to measure in a production setting due to the difficulty associated with standardising its assessment without using a standard human approach test which requires two experimenters, and a purpose-built arena in which to perform the test. Clearly this form of fear assessment is highly impractical for on-site welfare benchmarking, and consequently the development and validation of a fear of humans test able to be employed on-site across all production systems is recommended.

Before the on-farm pig welfare benchmarking protocol can be used in the Australian Pork Industry, a standardised training package needs to be developed to train pig producers in the effective on-farm application of its animal-based welfare indices. In order to combat the subjective nature of animal-based welfare indices, the scientific literature repeatedly reports that on-farm animal welfare assessment needs to be conducted by observers which are experienced and well-trained (Bokkers et al., 2009; Brenninkmeyer et al., 2007; D'Eath, 2012; Rutherford et al., 2009). Thus, in order to be effective an on-farm pig welfare benchmarking protocol, which is comprised predominately of animal-based welfare indices, requires a comprehensive training package to train operators in the effective on-farm application of its measures. Furthermore, given that pig producers will complete the training independently (i.e. without a trainer) the training package needs to be standardised in order to reduce the subjectivity and increase the feasibility of the on-farm pig welfare benchmarking protocol. This training package may be incorporated into a proposed project which will address stockperson training and assessment and welfare assessment for the pork value chain. In addition, we would suggest that APL survey target audiences and capture take-up, and maintain a database to measure improvements against base-line.

These recommendations are of particular importance if the proposed on-farm pig welfare benchmarking protocol is to be used effectively to benchmark pig welfare within the industry and to allow the collected data to be used to demonstrate welfare improvements over time in the Australian Pork Industry.

7. Intellectual Property

Information generated at this stage of the RD&E process, while creating intellectual property value, does not lead to patentable outcomes.

8. Technical Summary

The outcomes from the current project include;

- The modified BCS, LS and IS methodologies.
- Information regarding the validity and intra- and inter-observer reliability of the modified animal-based welfare indices (BCS, LS, and IS) for on-farm assessment in group-housed pigs

9. References

- Bokkers, E. A. M., Leruste, H., Heutinck, L. F. M., Wolthuis-Fillerup, M., Van Der Werf, J. T. N., Lensink, B. J., & Van Reenen, C. G. (2009). Inter-observer and test-retest reliability of on-farm behavioural observations in veal calves. *Animal Welfare*, 18(4), 381-390.
- Brenninkmeyer, C., Dippel, S., March, S., Brinkmann, J., Winckler, C., & Knierim, U. (2007). Reliability of a subjective lameness scoring system for dairy cows. *Animal Welfare*, 16(2), 127-129.
- Broom, D. M. (1986). Indicators of poor welfare. *British veterinary journal*, 142(6), 524-526.
Broom, 1998
- Barnett, J. L., & Hemsworth, P. H. (2003). Science and its application in assessing the welfare of laying hens in the egg industry. *Australian Veterinary Journal*, 81(10), 615-624.
- Cohen, J. W. (1988) *Statistical power analysis for the behavioural sciences* (2nd Ed). Hillsdale, NJ. Laurence Erlbaum Associates.
- D'Eath, R. B. (2012). Repeated locomotion scoring of a sow herd to measure lameness: consistency over time, the effect of sow characteristics and inter-observer reliability. *Animal Welfare*, 21(2), 219-231.
- De Koning, R. (1985) *On the well-being of dry sows*. Doctoral thesis, University of Utrecht.
- Fraser, D. (1995). Science, values and animal welfare: exploring the 'inextricable connection'. *Animal welfare*, 4(2), 103-117.
- Hemsworth, P. H., Brand, A. & Willems, P. J. (1981) The behavioural response of sows to the presence of human beings and their productivity. *Livestock Production Science*. 8, 67-74.
- Kaler, J., Wassink, G. J., & Green, L. E. (2009). The inter-and intra-observer reliability of a locomotion scoring scale for sheep. *The Veterinary Journal*, 180(2), 189-194.
- Karlen, G. A., Hemsworth, P. H., Gonyou, H. W., Fabrega, E., David Strom, A., & Smits, R. J. (2007). The welfare of gestating sows in conventional stalls and large groups on deep litter. *Applied Animal Behaviour Science*, 105(1), 87-101.
- Knierim, U., & Winckler, C. (2009). On-farm welfare assessment in cattle: validity, reliability and feasibility issues and future perspectives with special regard to the Welfare Quality® approach. *Animal Welfare*, 18(4), 451-458.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.
- Main, D. C. J., Clegg, J., Spatz, A., & Green, L. E. (2000). Repeatability of a lameness scoring system for finishing pigs. *Veterinary Record*, 147(20), 574-576.

Martin, P., & Bateson, P. P. G. (1993). *Measuring behaviour: an introductory guide*. Cambridge University Press.

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

O'Callaghan, K. A., Cripps, P. J., & Murray, R. D. (2002). 28. Behavioural indicators of pain associated with lameness in dairy cattle. *Research in Veterinary Science*, 72, 10.

Patience, J. F. & Thacker, P. A. (1989). *Swine Nutrition Guide*. Prairie Swine Centre, University of Saskatchewan, Saskatoon. 149–17.

Rutherford, K., Langford, F. M., Jack, M. C., Sherwood, L., Lawrence, A. B., & Haskell, M. J. (2009). Lameness prevalence and risk factors in organic and non-organic dairy herds in the United Kingdom. *The Veterinary Journal*, 180(1), 95-105.

Smulders, D., Verbeke, G., Mormède, P., & Geers, R. (2006). Validation of a behavioral observation tool to assess pig welfare. *Physiology & behavior*, 89(3), 438-447.

Whitham, J. C., and Wielebnowski, N. (2009) Animal-based welfare monitoring: using keeper ratings as an assessment tool. *Zoo Biology*. 28, 545–560.

Winckler, C., Capdeville, J., Gebresenbet, G., Hörning, B., Roiha, U., Tosi, M., & Waiblinger, S. (2003). Selection of parameters for on-farm welfare-assessment protocols in cattle and buffalo. *Animal Welfare*, 12(4), 619-624.