SHEEP CRC

Unpublished Report

Document ID:	SheepCRC_1_26
Title:	Quantification of temperament in sheep
Author:	Horton, B.; Pirlot, K.
Key words:	sheep; temperament; measurement; electronic weighing

This report was prepared as part of the Sheep CRC Program 2007-2014. It is not a refereed publication. If the report is quoted it should be cited as:

Sheep CRC Report 1_26

Quantification of temperament in sheep

CRC report March 2008

Brian Horton and Karen Pirlot
Extensive Agriculture Branch, DPIW, Tasmania

Janelle Hocking-Edwards and Rachael Gould
SARDI, South Australia

Contents

Summary 4
Project Aims 6
Background 6
Previous work 6
Potential use of standard weighing equipment 7
Feasibility studies 7
Measurements recorded 9
Layout of this report 9
Scales 10
Gate switches 10
Flight speed 10
Tag reader 10
Visual behaviour scores (SA method) 10
Settling time 10
Vocalisations 11
Vocalisation score 11
Behaviour 11
Behaviour category 11
Computer 11
Synchronisation of data 11
Operators 12
Sheep 12
Automated behaviour scores - CV and median deviation 13
Repeatability 14
Young ewes and wethers. 14
Older Ewes 14
Young Lambs 15
Production traits 17
Maternal behaviour 17
Lamb survival 17
Weight and growth rate 18
Time to shear sheep 19
Information Nucleus Flock 20
Lamb and Ewe differences 20
Correlations 20
Sire effects 21
Factors affecting lamb weight 22
Conclusions 23
Variations in method 24
Minimum weight difference 24
Optimum time period to record weights 24
Weights removed at the start of the weighing period. 24
Weights removed at the end of the weighing period 25
Conclusions 26
Dual weight recording units 26
Pairing of front and back weights 29
Comparison of single and dual weigh units on the same day 29
Ewes and Lambs with dual scale readout 31

Flight speed....32
Factors affecting temperament measurement....32
Previous experience....32
Other factors....32
Comparison of CV of weight with visual scores....33
Conclusions....34
Further work....35
Pregnancy scanning and lamb survival....35
Sheep familiarity with the weighing system....35
Other behavioural traits....35
Other production traits....35
Factors affecting the measurement....35
Optimum age to test the sheep....35
Standard protocol....36
References....36

Summary

Sheep temperament has been reported to be associated with a range of important production traits including mothering ability and lamb survival, mating behaviour, milk quality, growth and meat characteristics. In almost all cases quieter sheep have better performance than sheep with more nervous sheep. Most previous measures of sheep temperament are time-consuming or expensive.

A measure of sheep temperament may be obtained from the behaviour of the sheep while being weighed. Several methods are being compared. This report considers the methodology required for reliable temperament measurements and some production data related to these temperament scores.

Behaviour recorded by an observer. This is the lowest cost method and is reasonably repeatable but it is subjective, may require observer training and there may be significant operator differences. Various behaviour scores are being examined.

Settling time. This is the time required for the sheep to "settle" in the weighing crate with all feet and the head steady. This is also subjective, but easier to quantify and more repeatable than behaviour categories and it may be possible to measure it using only the scale.

*Movement of sheep on the s*cale. The weighing unit records multiple weights per second, which are used to obtain an average weight, but also provide a measure of the movement of the animal while in the weighing crate. Various measurements are being compared for the best measure of movement and the time period to hold the sheep for measurement.

Flight speed when the sheep leaves the crate. This can be measured using two or more detector beams across the race although other methods may be examined. This requires extra equipment, but does not require holding the sheep in the scale longer than the usual weighing time. This measure is repeatable but is not correlated with the other measurements.

Consistent results have been obtained with several different groups of sheep covering ages from 3 months to mature adults. Some of these variables have a repeatability of up to 0.62. There are significant differences between sire groups, indicating a heritable effect.

There are indications that the calmest sheep have higher growth rate and body weight, better maternal behaviour and require less time to shear.

Factors affecting the reliability of the measurement are being considered. The results may be affected by early experience, changes in the recording environment, sheep or people in front of the scale and random noises during weighing. These factors can usually be controlled so that measurements within a single run are comparable.

The repeatability of some of these measurements appears to increase with age of the sheep or possibly with greater experience at being weighed. Results at 6 to 7 months were highly repeatable and lambs scores at 7 months were correlated with their dam’s score.

Measurements have been recorded on the Information Nucleus Flock at Struan, South Australia. These results have shown relationships between the scale movement method and visual scores for behaviour and settling time. The results will be used to investigate the heritability of behaviour and its relationship to other production traits.

Project Aims

Develop a measure of the repeatability of quantifying temperament in sheep using standard weighing equipment. By December 2009, document a standard protocol for reliable measurement of temperament in sheep using standard weighing equipment. The document will include the relationship of the measure to other important sheep traits. Establish proof of concept for quantification of temperament in sheep.

This is an interim report to review the methodology being tested and consider some of the biological consequences of temperament on sheep production.

Background

Previous work

Murphy *et al* (1994) studied several measures of ewe temperament and showed that some are related to maternal behaviour. Most of these were rather time-consuming to measure, particularly the arena test, in which Kilgour and Szanatr-Coddington (1995) and Kilgour (1998) found that ewes and rams selected for superior mothering ability had lower movement scores.

Blache & Ferguson, (2005) developed a method of measuring sheep temperament by holding the sheep in a fully enclosed box for a short period. Results using this method have been reported to relate to ewe lambing percentage, possibly because ewes that are easily frightened may be more likely to desert their lambs (Murphy *et al*, 1994). However, calm ewes were found to be more receptive to rams (Gelez et al, 2003) so the effect may not necessarily be due to maternal behaviour alone. Sart (2005) found that calm ewes had higher milk quality (higher protein) so there may be several reasons for the improved fertility in calm ewes. The arena test results are correlated with the results from the enclosed box method (Beausoleil *et al*, 2008).

The enclosed box method has a high repeatability, but requires a relatively specialised system and it is possible that an approximation of the same method might be obtained using an ordinary sheep weighing unit. Starbuck *et al*, (2006) used a standard weighing crate and recorded the time required for the sheep to become stationary (settling time). They reported a correlation of 0.4 between successive readings on the same sheep. This suggested that methods used during normal weighing would be satisfactory without expensive equipment.

The South Australian group studying settling time and other visual scores of behaviour is preparing a separate detailed report, which is summarised briefly here. The correlation between repeated measurements of settling time and flight time increases with age. Sheep with short settling times (i.e. quieter sheep) have higher wean weight and lower fat depth. Lambs with long settling times had lower loin glycogen and shear force and lower loin pH 24hrs after slaughter. Crossbred lambs classified visually as calm had heavier liveweights, greater carcass weights and fat depth and more tender loins with lower pH than lambs classified as nervous or ‘escape’.

Potential use of standard weighing equipment

Normally the weighing system reports an average weight after it has obtained sufficient readings to provide a steady result. However, the weighing units are capable of sending continuous readings to a computer so that the variation and movement of the sheep in the weighing crate can be measured. This can also record the time to become stationary as above.

If the weighing system can be shown to be suitable for assessment of temperament, the weighing units could contain additional programs, for example to provide a coefficient of variation (CV) of the weight during the weighing period, rather than only an average weight.

In this report 'Calm' sheep are those with low variation in weight, or low behaviour scores and 'Jumpy' sheep are those with high variation during weighing or high behaviour scores. High movement is usually taken to indicate fear (Forkman *et al*, 2007), although Beausoleil *et al*, (2008) suggested that sheep with low movement scores may actually be more fearful than sheep with high movement, so the term 'calm' refers to their movement, not necessarily their state of mind or personality.

Flight time has been reported to be useful for cattle (Fell *et al*, 1999). This is the time to cover a measured distance after the cattle leave the weighing crate. Those with short flight times (i.e. faster) had lower growth rates. It has been suggested that the method is not suitable for sheep (Blache and Ferguson, 2005), but it is being recorded in these studies, since it can be measured without affecting the other results.

Feasibility studies

In a previous report, Horton & Pirlot (2007) showed that either Ruddweigh scales, with 2 readings/second or Tru-test scales, with 10 weights/second could be used to measure the variation of weight during normal weighing. For example, the coefficient of variation of all weights recorded during the weighing period had a correlation of 0.4 to 0.6 between successive measures of the same sheep, at 2 to 4 week intervals.

The method was improved by the use of switches on the weighing crate gates. These indicate whether the gates are open or closed and simplify the data analysis because it is clear when the animal is securely inside the weighing crate. This also allows some extra measurements of the speed of entry of the sheep into the crate and the speed of exit. It is expected that a method will be developed that does not require the switches for commercial use, although some weighing systems do control gates automatically.

Examples of results are shown in Figure 1, with the trace of actual weights recorded, with markers showing the time when the back gate closed with the sheep in the crate, and the time when the front gate opened to release the sheep.

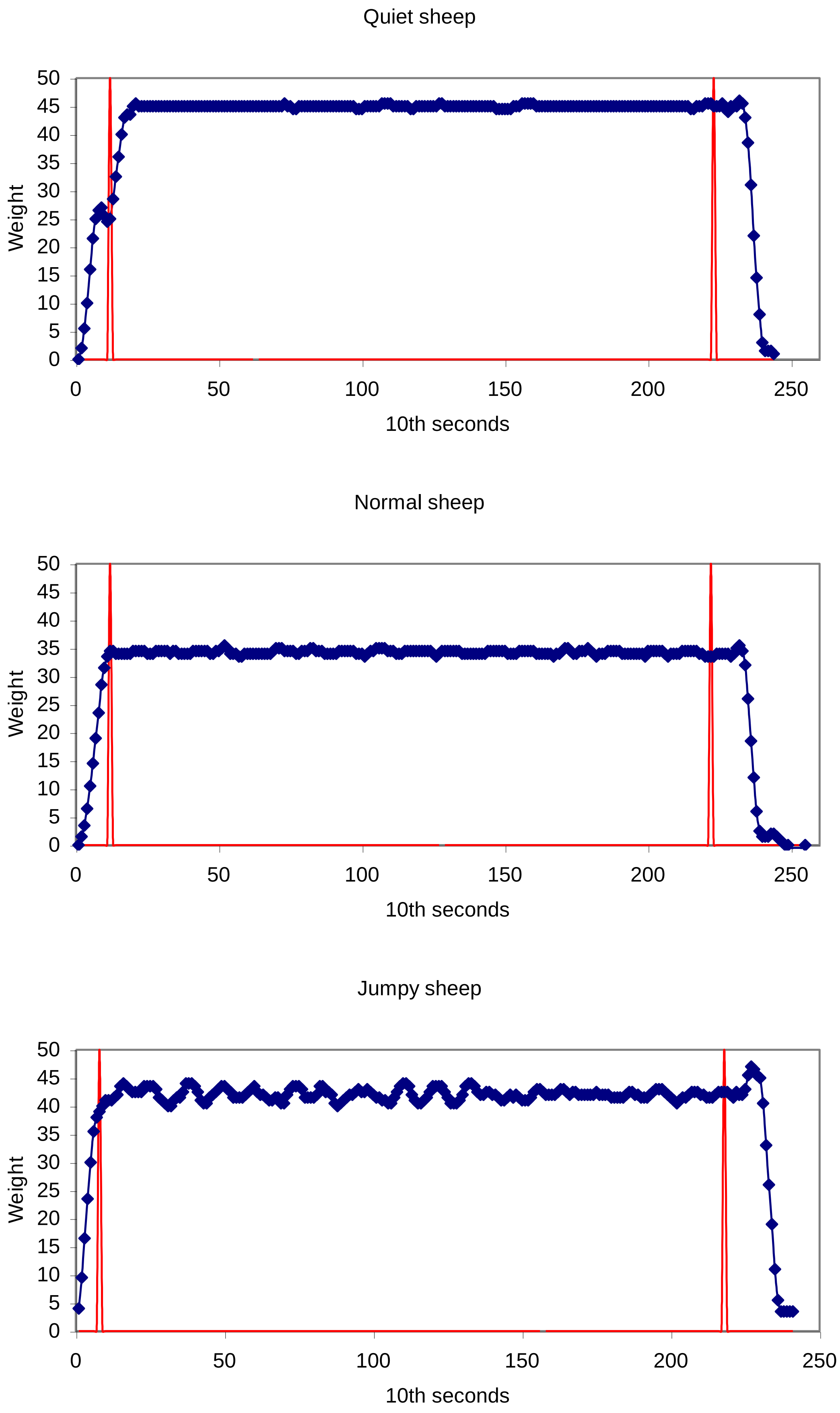

Figure 1. Example weight profiles showing gates shut and open.

Measurements recorded

The previous report (Horton & Pirlot, 2007) compared a wide range of measures of behaviour that could be obtained from the scale data or by direct observation. The measures with the highest repeatability were the CV of the weight and the median deviation of the weight. Some other values with high repeatability could be used but are more difficult to calculate during weighing (e.g. the intercept of the line fitted to the average absolute deviations) or depend on arbitrary values (e.g. the percentage of "low" deviations from the average weight).

The median of the deviation of weights from average was used because it was noted that a few sheep had relatively high CVs due to short periods of active movement that were not representative of the whole period. Use of the median ignores these short periods almost completely. However, the CV is a more common measure of variation and is easier to calculate.

The 'weighing period' is the time when both gates are closed, excluding the first 20 and last 10 readings. The number of records that should be exclude will be considered later in this report

Layout of this report

Most of this report relates to variations of the methods, so the report is not set out as the usual Methods, Results and Discussion. Instead the various trials are grouped under relevant subject headings.

The Methods section describes the equipment required, the sheep used and the scoring system for visual methods.

The first results section considers the repeatability of the method for different groups of sheep, usually with trials about 2 weeks apart, but in some cases with studies in the morning and afternoon of the same day to test potential new methods.

The next section considers the relationship of temperament scores to some production traits.

The results for the Struan Information Nucleus Flock are reported, with a comparison of the visual methods usually used there with the automated methods used in Tasmania.

Finally some variations on the standard method are reported, testing the optimum method of obtaining repeatable results and testing factors that disturb the temperament score. The possible use of dual scale read-outs is included here as a potential variation of the standard method. Measurement of Flight Speed is also reported in this section since the tests were run at the same time as the dual readout series and a home-made flight time system was being tested.

Methods

Scales

A Tru-test scale readout was set to send readings continuously to a computer. For some trials separate readout units were connected to the front and back load cells to send these results independently, but for most work both load cells were connected to the same readout and averaged.

Gate switches

Magnetic switches were attached to the front and rear gates of the weighing crate so that the switches were closed when the gates were closed. Each switch was connected to a relay box that converted the continuous on or off signal to a brief signal each time a gate opened or closed. These short signals were sent to a modified keyboard that interpreted the signal as a specific key being pressed on the numeric keypad. This allowed the computer to keep track of the position of each of the gates at all times. If the gate switches malfunctioned they could be simulated by pressing the required keys manually.

Flight speed

Temperament studies in South Australia used a Ruddweigh flight time recorder. This had two beams and reflectors. The first was placed just in front of the weighing crate, and the other 1.67m further down the race. The Ruddweigh recorder measures the time from when the first beam is cut to the second beam and displays the result in seconds. The time can be sent by serial link to a computer. This had a maximum time of about 6 seconds.

In Tasmania a home made system was used with three beams and reflectors. These sent a signal to a black box that started a timer when the first beam was cut, then measured the time to when the second and third beams were cut. As soon as the third beam was cut it sent both times through a serial link to the computer. This had a maximum time of 9 seconds. Due to the construction of the race it was difficult to place a beam at 1.67m so beams were positioned just beyond the crate and at 1m and 2m. Results reported here are usually the 2m time.

Tag reader

A panel reader was usually used to record electronic tags. The Tasmanian panel reader occasionally had power supply problems so a hand held wand reader was sometimes used. Waving the wand reader near the sheep's head sometimes caused extra jumpiness, so the panel reader was always used if possible.

Visual behaviour scores (SA method)

Settling time

The animal enters the crate, the rear door is closed and the time that it takes for the animal to become stationary is recorded. Stationary is defined as feet not moving and no significant movement of the head or body for a period of 5 seconds. (Previously reported Tasmanian results used a much less strict definition of settling time and are not included in this report).

Vocalisations

The person taking the recordings counts the number of calls that the animal makes during the first 30 seconds in the crate.

Vocalisation score

The animal is given a separate score for vocalisation. This score is 0 for no vocalisation, 1 for slight vocalisation (once or twice), 2 for moderate vocalisation (3 to 5), 3 for major vocalisation (persistent vocalisation).

Behaviour

The animal is given a behaviour score. This score is 1 for basically no movement and settling promptly, 2 for a little movement, 3 for more agitated movement and 4 for lots of extreme movement.

Behaviour category

The behaviour of the animal while in the crate is also categorised as "calm", "escape", or "nervous".

Computer

A portable computer was used to record data. This had serial connections to (a) the panel reader, (b) a wand reader if required, (c) a tru-test readout, (d) a second Tru-test readout if required, (e) the black box for flight speed if used. The computer had an external modified keyboard connected to the relay box receiving signals from the gate switches.

The Ruddweigh readout and behaviour scores were entered manually, with settling time recorded by hitting the space bar, when the sheep had settled. The computer recorded the order in which the sheep were weighed and the time of weighing as seconds after midnight.

On some occasions there were more serial connections than the computer could accept, so a second computer was used to record tag numbers or other information that did not require recording to the nearest millisecond. The second computer was synchronised to the nearest second to allow later merging of records by time of input, or the two computers were connected by yet another serial connection.

Synchronisation of data

On some occasions the panel reader, (particularly the SA version which forms a complete loop around the sheep), was able to read the entering sheep's electronic tag while only its head had entered the crate, before any weight was applied to the scale. On other occasions the panel reader, (particularly the Tasmanian one), did not read the tag until the weighing period was complete. The gate operator was often able to allow the incoming sheep to enter while both gates were open and before the weight of the outgoing sheep had fallen close to zero. Therefore certain rules were required to relate tags to sheep records.

The program required the tag number to be read between the time when (a) the weight of the *previous* sheep had dropped by at least two thirds and (b) the time when the gate was opened to release the sheep *after* weighing. The last tag read during that period was allocated to that weight record.

The weight variation was recorded from the 15th reading after both gates closed with the sheep in the crate to the 10th reading before the gate opened to release the sheep. For most records a minimum of 180 readings (18 seconds) was required for a valid result, or 300 readings for SA data, where sheep were held for a minimum of 30 seconds.

If the gates closed without the sheep's legs fully inside the crate, or with another sheep's head in the gate, the gate could be reopened and closed and as long as the weight did not fall to one third the previous average. In these cases the program assumed the same sheep was still inside the crate, but restarted the timer for weight records.

Operators

Apart from visual behaviour scores, the system recorded data automatically although an operator was required to open and close the weighing crate gates. A second operator moved the sheep up from behind to maintain a continuous flow. This was done without dogs and with no shouting or loud noises, to avoid startling the sheep being weighed.

Sheep

The sheep studied in South Australia (for this report) were the CRC Information Nucleus Flock at Struan. The 2007 drop lambs and their mothers were tested in November 2007. The results for this group are reported in the section **Information Nucleus Flock**.

The sheep tested in Tasmania, and used for all the other studies in this report, were part of the flock at Cressy Research Station.

The 2005 drop ewes and wethers, included about 200 of each gender. These were studied four times under standard condition from April to July 2007 and again using a variety of methods after the ewes' lambs were weaned.

A group of 300 'older ewes' were also tested twice before lambing in 2005. These had all lambed at least once in previous years and had all been pregnancy scanned as having at least one lamb in 2007. Half of these ewes were included in a lambing study requiring blood samples from the lambs before suckling, giving a measure of the ewes' behaviour during sampling and a record of lamb survival for each ewe. These ewes in the lambing study were tested again immediately before weaning.

The 2007 lamb drop were tested several times, but with different groups each time. The lambs born to the ewes in the lambing study were tested, with their mothers, immediately before weaning, then all 480 surviving lambs born to the 300 older ewes and the 200 2005 drop ewes were tested twice after being combined in a single mob. This included the 'lambing trial' group, which were combined with all other lambs by about 5 months of age. Finally about 130 of these lambs were tested after the heaviest lambs had been sold and some others removed for other unrelated trials. They are not representative of the entire lamb drop but were used for repeatability studies.

Automated behaviour scores - CV and median deviation

Apart from visual scores and flight times, the primary method used was the coefficient of variation (CV) of the weight during the weighing period. This is the standard deviation as a % of the weight, so should be independent of the body weight.

An alternative measure is the median deviation. This is obtained by determining the mean weight, expressing each weight record in the weighing period as the absolute percentage deviation from the mean weight, then taking the median of those values. This is more complex to calculate but is less affected by brief random movements.

Repeatability

Young ewes and wethers.

The Cressy Research Station 2005 drop ewes and wethers (about 200 of each) were recorded on four occasions at intervals of 2 to 4 weeks at about 18 to 22 months of age. The correlations between the successive results are shown in Table 1.

Table 1. Correlations of ewe and wether results between each weighing trial

Variable	1v2	2v3	3v4	Ave1&2 vs Ave3&4
CV Ewes	0.54	0.49	0.46	0.62
MedianDev Ewes	0.57	0.49	0.46	0.57
CV Wethers	0.36	0.52	0.45	0.53
MedianDev Wethers	0.57	0.55	0.45	0.60

The CV and median deviation provided good consistency between each weighing trial.

In the case of the wethers, for management reasons it was necessary to conduct the fourth trial in a different shed, which the wethers were not familiar with. In addition this shed had a much less efficient system for moving sheep into the race leading to the scales. Despite these differences there was still a significant correlation with previous results. Therefore the difference in environment did not prevent reproducible measurements.

There is a very high correlation between CV and Median deviation (0.86), indicating that they are simply different methods of measuring the variation in movement of the sheep during weighing. CV is easier to measure and covers the entire period, but median deviation is less subject to brief random movements.

For the measurement with the best overall correlations, median deviation, the repeatability is 0.62.

Older Ewes

A group of 300 older ewes (3 to 6yo) were tested twice between mating and lambing. These ewes had all been pregnancy scanned and half of them were used for a study on lamb survival, which required observation during lambing.

The average score of these older ewes (Table 2) for CV and Median deviation was about 75% that of the younger ewes. These ewes had been weighed on many previous occasions in previous years and in the period before and during mating they were weighed and condition scored at two week intervals.

Despite their frequent exposure to the weighing system and their lower results, there was still variation within the group and the correlation between the two tests before lambing was as high as for the young ewes. This indicates that repeated exposure to the weighing system does not prevent the recording of repeatable results. It may be that some exposure to the weighing system may be beneficial in reducing random effects for some sheep.

These ewes were tested again on the day the lambs were weaned, but before weaning, with the lambs and ewes mixed in the race. The agitation due to the mix of lambs and ewes weighing may have contributed to a lower correlation compared with the previous results. However, the scores are the same as obtained before lambing and the results are still very consistent. Although testing ewes with lambs at foot during testing is not recommended the results suggest that the temperament score is reasonably robust as a consistent measure.

Table 2. CV and Median deviation for young ewes and wethers and older ewes

Variable	2005 drop wethers	2005 drop ewes	Older ewes trial 1	Older ewes trial 2	Older ewes weaning
CV	1.43	1.34	0.97	0.94	0.97
MedianDev	0.96	0.91	0.68	0.68	0.66

Table 3. Correlation of repeated results for older ewes

Variable	Trial 1 vs Trial 2	Trial 1 vs weaning	Trial 2 vs weaning
CV	0.62	0.52	0.58
MedianDev	0.69	0.44	0.48

Young Lambs

The 2007 drop lambs were tested at weaning (3months), and again at 5, 6 and 7 months of age. As described above, the test before weaning was carried out without drafting lambs from their mothers. The lambs and ewes referred to above formed part of a lambing trial and these sheep were weighed regularly from before weaning to 7 months.

Table 4. Mean and standard error of CV for lambs

Age (months)	Lambing Trial	Other Lambs
3 (weaning)	1.76 ± 0.06	Not tested
5	1.25 ± 0.05	1.45 ± 0.04
6	1.34 ± 0.05	1.83 ± 0.04
7	1.06 ± 0.06	1.57 ± 0.05

The two groups are significantly different on all occasions tested (p<1%)

The other lambs that were not part of the lambing trial were not weighed until they were combined with the lambing trial group just before 5 months of age, after which they were kept with that group and weighed together. These other lambs had less early experience of being weighed and had higher scores at every occasion when they were tested. Although they were all being weighed together every 2 to 3 weeks after 5 months the difference between the lambing trial group and the others increased and became more significant over time. This indicates that it is important only to compare sheep that have similar experience at being handled and weighed. Therefore the method would not be suitable to compare animals purchased from different sources, but this is applicable to many other production measurements.

The correlation of the early results with later results is poor, until 6 months of age, after which a high correlation is obtained with later results. There was a very high correlation between tests on the same day at 7 months.

Table 5 Correlation of early CV with later CV for lambs

Age	5m	6m	7m AM	7m PM
3m	0.15	0.16	0.17	0.49
5m		0.26	0.22	0.23
6m			0.50	0.56
7m AM				0.65

The 7 month trial was carried out in the morning and again on the same sheep in the afternoon.

The correlation of lamb results with their mothers CV was low until the 7 month results, when it became significant (0.46). These results suggest that testing lambs at a very early age may not provide reliable results, but results at 6 months are sufficiently consistent.

Production traits

Maternal behaviour

Preliminary studies (Horton and Pirlot 2007) had suggested that maiden ewes with lower CV had a higher survival rate for single lambs. However, there was a limited number of ewes available for the study and lamb deaths can be due to a wide variety of causes.

In 2007 the older ewes were part of a lamb survival study that required blood samples from the lambs before suckling. This gave the opportunity to measure the distance the ewes moved from the lamb during sampling and the time taken to return to the lamb afterwards.

Table 6 shows that the calmer ewes (CV <=0.75) moved a shorter distance from the lamb and returned within a shorter time than the ewes with higher CVs. (p <5%)

Table 6. Difference in ewe behaviour between calm ewes (CV <= 0.75) and jumpy ewes (CV > 0.75). Mean and SE of distance (m) the ewe moved during lamb bleeding and minutes taken to return to the lamb.

Ewe behaviour	Calm ewes	Jumpy Ewes	t
Time to return (min)	1.00 ± 0.32	7.16 ± 2.97	2.07
Distance moved (m)	15.4 ± 2.1	24.4 ± 3.2	2.34

Lamb survival

The ewes with dead lambs had slightly higher CV than those with no dead lambs (Table 7). However the only significant effect (p < 5%) was the higher Median Deviation for ewes with lambs that died more than 24 hours after birth but before weaning.

This data is currently being analysed further to differentiate effects of lamb birth weights and weather, which tend to dominate other factors affecting lamb survival.

Table 7. Mean and SE for CV and Median Deviation of ewes differing in survival of their lambs.

Lamb survival	CV	Median deviation
Survived to weaning	0.904 + 0.036	0.642 + 0.028
Died within 24hrs	0.873 + 0.047	0.636 + 0.033
Died after 24 hrs	1.053 + 0.079	0.780 + 0.054*
All Deaths	0.955 + 0.045	0.702 + 0.032

* significantly different from ‘Survived to weaning’, p<5%

Weight and growth rate

The 2005 drop ewes had been weighed several times before and during testing so body weight and growth rates could be compared with temperament scores.

The jumpier sheep had significantly lower body weight from five months of age to 18 months (Table 8). During periods when growth rate was high the jumpier sheep had a significantly lower growth rate (Table 9). However, when growth rate was low there was no association between temperament score and growth rate.

Table 8. Effect of 1 unit of median deviation on Body weight (by linear regression)

Variable	Factor	SE	t	p
Weight 10/1/06	-1.34	0.63	2.13	3.3
Weight 10/3/06	-1.70	0.60	2.82	0.5
Weight 12/5/06	-1.30	0.60	2.18	2.9
Weight Jan-May06	-1.45	0.58	2.48	1.3
Weight 15/10/06	-1.96	1.04	1.88	6.0
Weight 10/2/07	-2.30	1.09	2.11	3.5

Table 9. . Effect of 1 unit of median deviation on growth rate (by linear regression)

Variable	Factor	SE	t	p	Growth (g/day)
Growth Jan-Mar06	-6.0	5.1	1.17	NS	9.4
Growth Mar-May06	6.3	4.4	1.43	NS	5.5
Growth Jan-May06	0.36	2.92	0.12	NS	7.4
Growth May-Oct06	-11.2	4.9	2.28	2.2	50
Growth Oct-Jan07	-2.9	4.9	0.58	NS	79
Growth May-Jan07	-7.5	3.0	2.53	1.1	63

This is consistent with the results of Fell et al, (1999), who found that cattle with a high flight speed had lower growth rate than slower cattle.

Time to shear sheep

During shearing in 2007, the time taken to shear each individual sheep was measured for the 2005 drop ewes and wethers and the older ewes. Allowance was made for different times between shearers, effects of age and gender and differences between mulesed and unmulesed sheep. The most jumpy group of ewes (25% with highest CV) took 4% longer to shear than the calmest group (40% with lowest CV).

The shearers were learners, two of whom had been shearing only a few weeks. It is not clear whether more experienced shearers would have less difficulty handling even the most jumpy sheep (i.e. giving less effect of temperament) or whether there would have been more random events affecting these learner shearers. Therefore it is not certain whether these effects would apply to experienced shearers, although there does appear to be a measurable difference in ability to handle these sheep. This handling difficulty may apply to other situations such as drenching, crutching etc.

Information Nucleus Flock

The CRC Information Nucleus Flock at Struan was tested using both the automated measurement of CV of weight and using the visual behaviour scores. Most of the lambs were tested on one day, then the remaining 100 lambs and then the ewes were tested on the next day. Flight time was recorded with the Ruddweigh unit at 1.67m.

Lamb and Ewe differences

The lambs were lighter and had scores indicating more nervous and jumpy behaviour (Table 10).

Table 10. Ewe and lamb means and standard deviations for selected measurements and correlations between lamb scores and their mother's result

Measure	Lambs	Ewes	Correlation
Weight	30.6 ± 7.3	63.3 ± 9.0	0.34
Calls in 30sec	0.73 ± 1.60	0.3 ± 1.1	-0.01
Call Score	0.40 ± 0.73	0.16 ± 0.48	-0.04
Behaviour	2.46 ± 0.68	2.04 ± 0.85	0.13
Flight time	0.97 ± 0.42	1.30 ± 0.95	0.22
Settling	25.2 ± 12.4	17.0 ± 13.4	0.06
CV of weight	1.49 ± 0.63	1.06 ± 0.54	0.07
Median Deviation	0.90 ± 0.33	0.62 ± 0.35	0.11

Correlations

The measures of settling time, behaviour score, CV and median deviation were correlated with each other (Table 11), suggesting that they measure a similar characteristic.

Call scores and flight times were not correlated with the above behaviour scores, nor with each other.

Only flight speed was correlated with weight, and only in lambs, but this needs to be considered after allowing for sire affects, which strongly affect weight (see below).

The correlations between lamb and ewe scores were low, although they were positive for most of the important measures. The poor correlation may be due to the effect of sire differences, discussed below or the greater variation in very young lambs discussed elsewhere.

Table 11. Correlations between selected measures.
Ewes above the diagonal, lambs below the diagonal

	Wt	Calls	Call Score	Behav	Flight	Settle	CV	Med Dev
Wt		0.09	0.06	0.01	0.06	-0.02	-0.09	-0.13
Calls	-0.02		0.88	0.10	-0.08	0.08	0.03	0.03
CallSc	-0.01	0.94		0.15	-0.08	0.11	0.05	0.05
Behav	0.12	0.10	0.10			0.62	0.70	0.67
Flight	-0.25	-0.04	-0.04	-0.10		-0.10	-0.18	-0.15
Settle	0.08	-0.01	-0.01	0.45	-0.12		0.58	0.58
CV	-0.05	0.04	0.03	0.51	-0.18	0.33		0.94
MdDv	-0.10	-0.04	-0.05	0.52	-0.12	0.44	0.78	

The agreement between different behaviour measurements was generally better in the ewes than in the lambs, suggesting that there are more random factors affecting the scoring and recording of lamb behaviour.

Sire effects

Sire groups were Merino (M), Poll Merino (PM) and Border Leicester (BL) mated to Merino ewes and Hampshire Down (HD), Poll Dorset (PD), Suffolk (SF), Texel (TX) and White Suffolk (WS) mated to crossbred (XB) ewes.

There were very highly significant differences in weight between groups (Table 12). The Merino and Poll Merino Sire Groups were much lighter than the others at 23kg, then the Border Leicester at 30kg, with the cross bred ewe groups mated to terminal sires the highest weight (34-38kg).

The Merino, Poll Merino and White Suffolk had the lowest behaviour scores and CV but although the two Merino groups shared the slowest flight time the White Suffolk were much faster. The Suffolk and Texel had the highest behaviour scores and the Suffolk the shortest flight time.

It is not possible to separate effects of XB mothers from effects of sire groups because cross-bred ewes were mated to HD, PD, SF, TX and WS sires and Merino ewes were mated to M, PM and BL sires.

Table 12. Sire Group means and standard errors for selected values

Sire	Weight	Settling time	Behaviour	CV	Median Deviation	Flight time
M	24.3 ± 0.6	23.0 ± 1.6	2.28 ± 0.09	1.27 ± 0.06	0.79 ± 0.03	1.10 ± 0.08
PM	23.8 ± 0.04	23.2 ± 1.2	2.35 ± 0.07	1.50 ± 0.08	0.90 ± 0.03	1.11 ± 0.06
BL	30.4 ± 0.06	27.8 ± 1.7	2.56 ± 0.08	1.62 ± 0.09	0.89 ± 0.03	0.93 ± 0.04
HD	38.9 ± 1.7	28.1 ± 3.4	2.50 ± 0.17	1.37 ± 0.12	0.82 ± 0.06	0.81 ± 0.08
PD	36.5 ± 0.6	25.2 ± 1.4	2.56 ± 0.07	1.50 ± 0.10	0.86 ± 0.04	0.91 ± 0.03
SF	37.5 ± 0.9	29.5 ± 4.3	2.83 ± 0.21	1.71 ± 0.27	0.86 ± 0.09	0.73 ± 0.05
TX	34.1 ± 1.0	30.7 ± 4.1	2.65 ± 0.17	1.69 ± 0.14	0.90 ± 0.06	0.86 ± 0.06
WS	36.8 ± 0.6	20.8 ± 1.0	2.38 ± 0.10	1.34 ± 0.70	0.74 ± 0.04	0.85 ± 0.04

Pairs of significantly different results are not indicated here, but all columns contain sire groups that are significantly different from some other sire groups.

Factors affecting lamb weight

Multiple linear regression was used to determine the factors most strongly related to lamb weight on the day temperament scores were carried out (Table 13). The weights were recorded about 2 weeks after weaning, and due to the use of AI, all lambs were born within a few days of each other. Therefore the weights are similar to wean weights and represent growth rate to weaning.

There were highly significant effects of sire breed on lamb weight as expected and these accounted for most of the variation in body weight (r^2 = 0.610 when only sires included). However, differences in median deviation had a similar level of effect to some of the differences between sire breeds.

High median deviation scores were related to low body weight, after allowance has been made for sire breed. CV was also significant when median deviation was excluded.

Behaviour scores, flight speed, call scores and settling time were not significantly related to body weight.

Table 13. Multiple linear regression for lamb weight (r^2 = 0.62)

Factor	Coefficient	t	p (%)
Sire BL	6.20 ± 0.84	7.35	<0.1
Sire HD	14.71 ± 1.57	9.35	<0.1
Sire PD	12.35 ± 0.82	15.11	<0.1
Sire SF	13.40 ± 1.45	9.21	<0.1
Sire TX	10.03 ± 1.14	8.78	<0.1
Sire WS	12.34 ± 0.88	14.02	<0.1
Sire PM	-0.19 ± 0.82	0.82	n.s.
Day 2 (vs Day 1)	-1.32 ± 0.65	2.04	4.1
Order	-0.0060 ± 0.0027	2.21	2.7
Median deviation	-2.11 ± 0.79	2.66	0.8

Sires were compared with Merino sires
Order is the order in which sheep were weighed, starting from 1 on each day.
Factors not listed were not significant, or were weaker measures of the factors listed (e.g. CV). Dam breed was confounded with sire breed.

Conclusions

Visual and computer scores of behaviour traits were correlated, confirming that they are measures of the same behaviour. There were weak correlations between lamb's scores and their mother's scores, but there were significant sire differences, indicating a genetic effect. After allowing for sire effects on growth rate, some measures (e.g. median deviation and CV of weight) were inversely related to lamb weight (and therefore to growth rate to weaning).

Variations in method

Minimum weight difference

When weighing sheep the scale usually provides weights to the nearest 0.5kg. However, for sheep under 50kg it can be set to weigh to the nearest 0.2 kg. This affects the smallest differences between weights that can be measured allowing measurement of finer changes in movement.

For one series of measurements half the ewes were weighed with the scale set to record weights to the nearest 0.5kg and the others weighed to 0.2kg. Both methods gave good correlations with previous results, but the CV and median deviation were higher when the scale was set to 0.5kg and the variation in the CV (shown by its SD) was greater (Table 14). The 0.2 kg setting should be used if possible and has been used for most of the results reported here.

Table 14. Mean and standard deviation of CV for ewes and lambs weighed at the fine (0.2kg) or coarse (0.5kg) settings

Sheep	0.2kg	0.5kg
Ewes	1.35 + 2.00	1.64 + 2.57
Lambs	1.77 + 0.79	1.84 + 0.94

Optimum time period to record weights

In figure 1 it was clear that there is a lag period after the gates are closed and the sheep is fully inside the crate, before the weight reaches a stable level. Therefore the first few weights **after** the gate is closed must be discarded. There was also concern that the operator might affect the result **before** the front gate was opened when moving their hand to the gate lever and either adding weight to the scale or startling the sheep.

The data for test 2 and 3 of the combined ewes and wethers was used to examine the optimum period over which to record the weights. The correlation of CV between test 2 and test 3 was used as an indicator, with the aim of optimising this correlation.

Weights removed at the start of the weighing period.

Figure 2 shows that if all weights are included from the instant the gate is closed with the sheep inside, the correlation is not as good as when some of the early weights are removed. The highest correlation is obtained when about 20 weights are excluded, corresponding to the first 2 seconds after the sheep is fully inside the crate. Previous examination of weight charts has shown that it takes about 1.5 seconds for the weight to stabilise, but there may be slightly longer delays for some sheep.

If more weights are excluded the correlation begins to decrease, indicating that the loss of data is degrading the accuracy of the measurement.

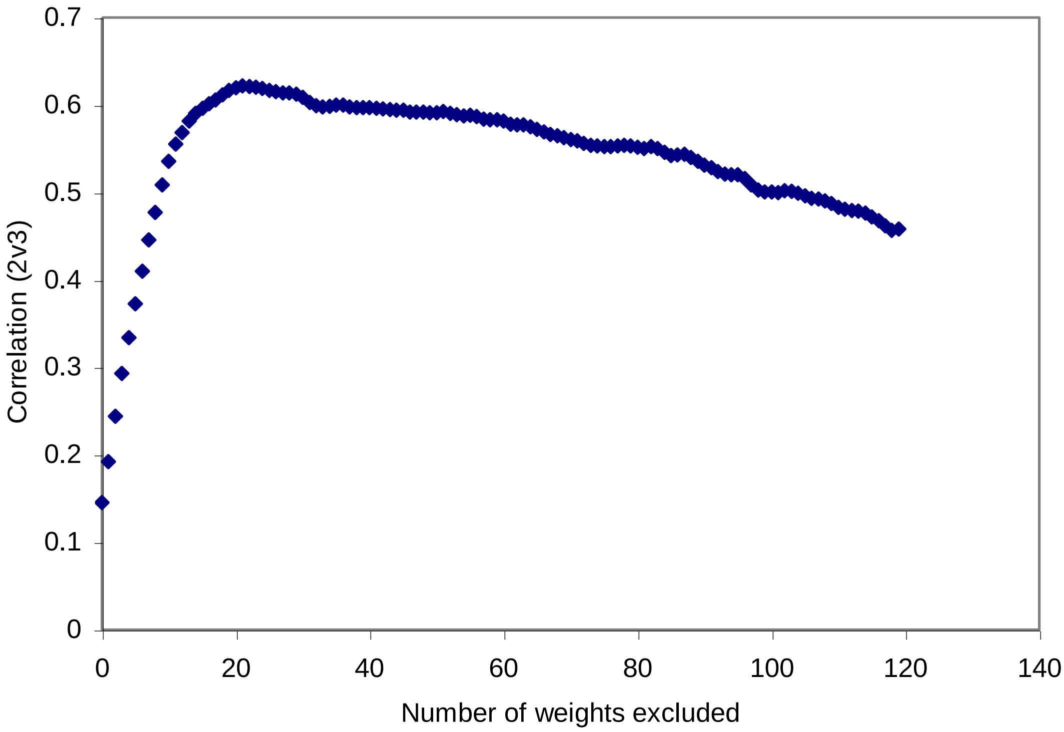

Figure 2. Weights excluded at the start of the weighing period

Weights removed at the end of the weighing period

In figure 2 the same process is applied but in this case the weights are excluded from the end of the weighing period (after excluding the first 20 weights). The results indicate that it is not essential to remove any weights before the gate is opened, suggesting that the operators hand on the gate or near the sheep's head does not cause any random fluctuation in weight before the gate is opened. Removal of weights from the end of the weighing period does not cause substantial lowering of the correlation unless about 50 or more weights are removed.

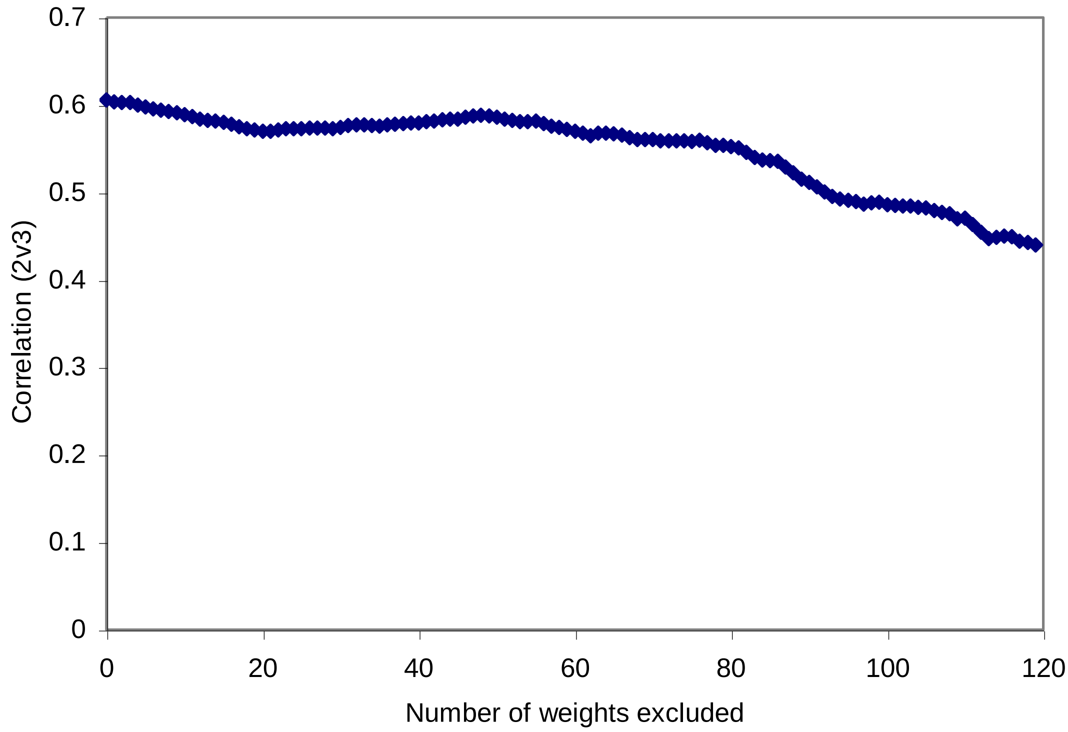

Figure 3 . Weights excluded at the end of the weighing period.

Conclusions

These results indicate that the first 20 weights should be excluded, but it is not necessary to exclude any at the end of the weighing period. They also show that more weight records improve the precision of the measurement. It is possible that weighing times longer than 20 seconds may be useful, but for most purposes about 15 seconds is adequate, while less than 10 seconds may be insufficient. In practice a balance is needed between incremental gains from additional time in the scale and efficiency in weighing sheep as quickly as possible.

Similar studies (not shown) using median deviation, found that it is less sensitive to removal of early weights, because to a limited extent the median ignores anomalies at the start (or end) of the weighing period. So only the first 10 weights need to be removed at the start when median deviation is used.

Dual weight recording units

The previous report suggested the use of separate readout units for the front and rear load cells. This may give a more informative result than merely averaging the two load cells. In practice this may not require two readout units, if the single unit is capable of separating the results from the two load cells.

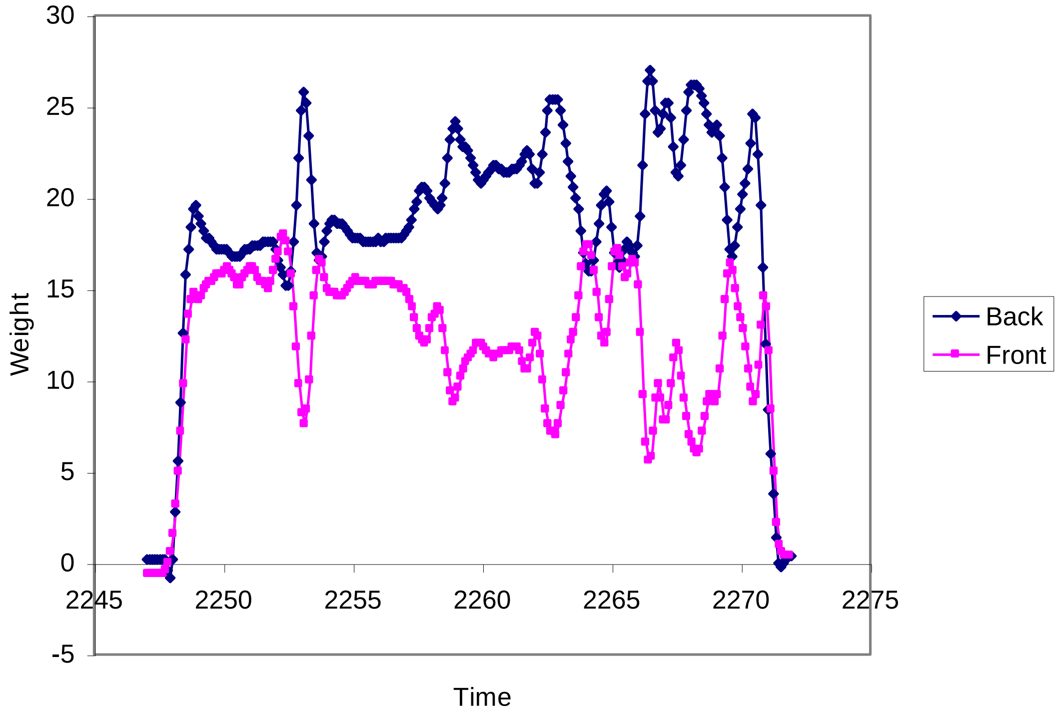

Figure 4. Wether 5501 tested using a separate readout for each load cell.

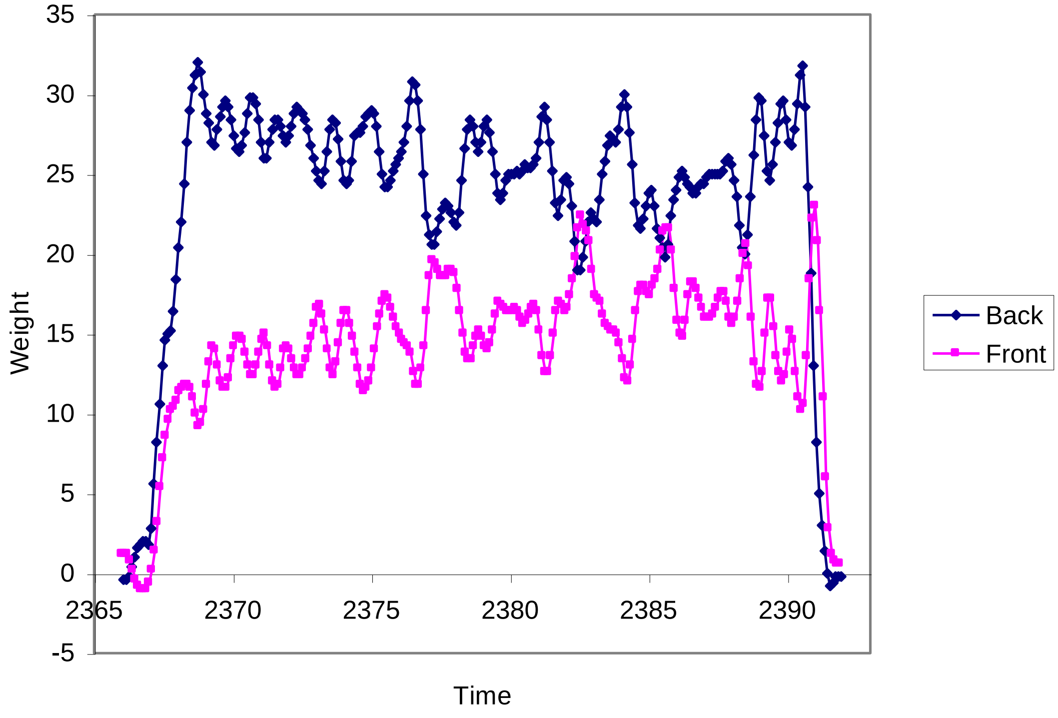

Figure 5. Wether 4462 tested using a separate readout for each load cell.

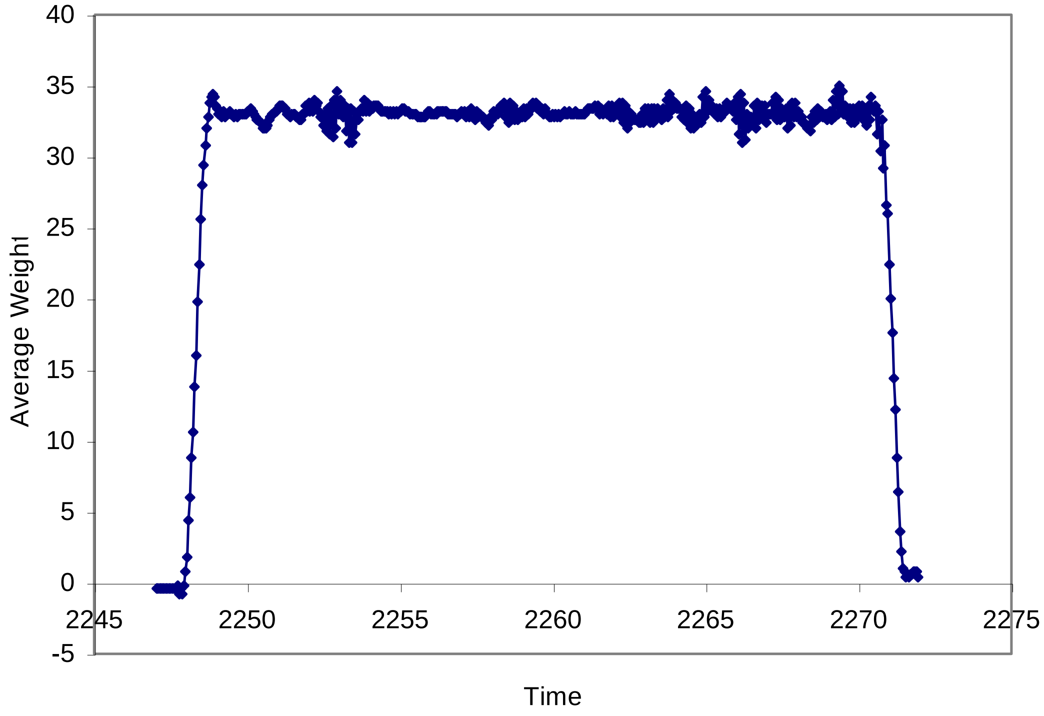

Figure 6. Wether 5501: Sum of separate readouts from each load cell.

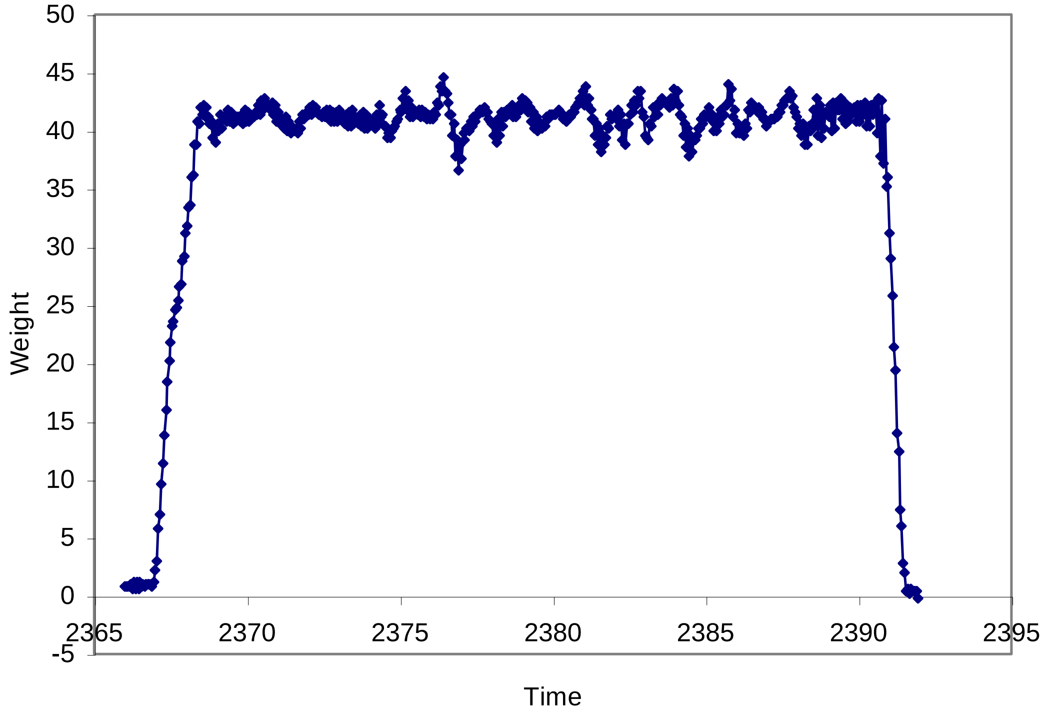

Figure 7. Wether 4462: Sum of separate readouts from each load cell.

Pairing of front and back weights

The weights from each load cells are recorded at slightly different times although less than 1/10 second apart. So for each weight on the front scale the two nearest weights before and after that time from the back scale were combined to give an estimated 'back' weight at exactly the same time as the front weight. A similar method was used to give estimated 'front' weights at exactly the same time as the back weights. This gave twice as many weights as usual, because each estimated weight was interpolated between the actual measured weights. However, the procedure was necessary in order to add the pairs of weights to estimate the weight in the usual way.

The 'total weight' is the sum of the actual weight and the corresponding estimated weight from the other unit.

'Weight difference' is the mean of (back weight minus the front weight) as a % of the total weight, giving a measure of the degree to which sheep consistently kept towards the back or front of the scale.

The SD of weight difference gives a measure of the degree to which the sheep moved backwards and forwards during weighing. (The CV from the total weight measures how much the sheep moves up and down).

Comparison of single and dual weigh units on the same day

The 2005 drop wethers were weighed and scored in the morning and again in the afternoon. On each occasion some were scored normally (single weigh unit) and others with double weigh units. This meant that over the two weighings some were weighed with a single unit on both occasions, some only with a double unit and other with each method. This allowed a measure of the consistency of each type of weighing and the relationship between the normal method and the dual-weigh-unit method

Table 15 shows the correlation between morning and afternoon measures of the same measurement on the same wether. Body weight and CV of body weight were considered to be the same measure whether done by single or dual scale method.

Table 16 shows correlations between different measurements after combining the morning and afternoon results.

There was a very high correlation (0.97 to 0.98) between weight in the morning and in the afternoon whether by single or dual scale method. This indicates that the links between records were made correctly and the system of adding weights from two weigh units is valid.

The 'normal' temperament scores were derived by adding the two weights and scoring variation as usual. This correlation between the two scores (0.62 to 0.70) was better than between tests a few weeks apart (usually 0.4 to 0.5) but consistent enough to indicate that the single and dual scale methods measure the same thing.

Some readings were only available using the dual scale method. The correlation between successive measures of weight difference was 0.72, indicating that sheep are consistent in the behaviour in staying towards the front or back of the weigh crate during weighing. The correlation of the SD of this value was only (0.38), suggesting that back and forth movement is less consistent than up and down movement.

Sheep with low CV (low up/down weight variation) tended to have more weight to the back, compared with those with high CV. However, there was no correlation between the tendency to keep consistently to the back (or front) and the variation in back and forth movement.

Although weight difference is consistent it is only moderately correlated with CV, suggesting that it is measuring something slightly different from the standard measure, which depends on up and down movement.

Table 15. Correlations between morning and afternoon results for the same wethers

Trait	Correlation
Single unit body weight	0.967
Double unit body weight	0.981
Single body weight vs double body weight	0.982
Single unit CV	0.695
Single CV vs Double CV	0.612
Double unit CV of total weight	0.623
Double unit Wt difference	0.717
Double unit SD of Wt difference	0.381

21 wethers by single method morning and afternoon, 22 by double method morning and afternoon, 55 by single then double, 15 by double then single

Table 16. Correlation between different measures of temperament

Traits	Correlation
Single CV vs Double Wt difference	-0.322
Single CV vs Double Wt Difference SD	-0.015
Double CV vs Double Wt Difference	-0.472
Double CV vs Double Wt difference SD	0.188
Double Wt Diff vs Double Wt Diff SD	-0.096

The dual-weigh-unit method obtained values with higher repeatability than the standard single unit temperament scoring method (CV of total weight), suggesting that it may be a more reliable measure of the trait that it is measuring.

Ewes and Lambs with dual scale readout

About 100 of the 2005 drop ewes were weighed using the dual readout system in the morning and again in the afternoon. On this occasion no measures were taken with a single readout, but this has been done at least 4 times before with these ewes. A similar study used 82 lambs from the 2007 drop (7 months).

The correlations for the various measures are shown in table 17 (ewes) and 18 (lambs). The weights were very consistent, as expected, confirming that the correct weight can be obtained by summing the front and rear readouts separately. The lamb weights had a weaker correlation because there was little variation in this group. The heavier lambs had been sold and the lighter lambs were in a separate group.

Table 17 Correlation of measurements on the same ewes in the morning and afternoon.

Measurement	Weight	CV	Wt diff	Diff SD	Time 2m
Weight	0.98	-0.17	-0.22	-0.28	0.05
CV		0.60	0.45	0.47	-0.24
% Wt diff			0.76	0.41	0.02
Wt Diff SD				0.54	0.00
Time 2m					0.71

Values **on** the diagonal are correlations between morning and afternoon. Values **above** the diagonal are between the combined measurements from the morning and afternoon.

Table 18 Correlation of measurements on the same lambs in the morning and afternoon.

Measurement	Weight	CV	Wt diff	Diff SD	Time 2m
Weight	0.85	0.08	0.08	-0.15	0.09
CV		0.59	0.52	0.59	-0.30
% Wt diff			0.79	0.41	-0.17
Wt Diff SD				0.58	-0.27
Time 2m					0.53

Values **on** the diagonal are correlations between morning and afternoon. Values **above** the diagonal are between the combined measurements from the morning and afternoon.

For both ewes and lambs there was a strong consistency in keeping the weight either towards the front or back and in the variation of front to back weight. These effects were about as strong as the usual CV of total weight and although they were correlated with this measure of temperament there may be other factors involved.

The double scale method is able to obtain temperament scores in the same way as usual, by using the sum of the front and back weights. However, it is also able to obtain additional measures of the degree to which the sheep stays near the back or the front of the weigh crate and the variation in this back and forth movement. Therefore some additional measurements should be recorded using this system, since it derives all the standard measures plus extra information without additional work during recording.

Flight speed

Flight speed has been measured routinely in South Australia and is reported separately for that site, where the repeatability is high. In the dual readout measure of the 2005 drop ewes and lambs (above) flight speed was also recorded using a beam just in front of the crate and other beams 1m and 2m from that beam (Tables 17 and 18).

The time over 2m was very consistent in the morning and afternoon readings, slightly more consistent than the times for 1m or for 1m to 2m (not shown). The time was only weakly correlated with the other behaviour scores, so although the flight time is consistent it measures a different aspect of behaviour. There was no consistency in the ratio of 1m to 2m times so the sheep are not consistently changing their speed over the distance measured.

It is clear that the single beam method at 1.67 or 1.7m should give good results although other distances and conditions may be tested.

The movement of sheep out of the weigh crate may be dependent on the type and length of race beyond the crate and whether other sheep are visible or not. This may be why Blache and Ferguson (2005) reported extremely low repeatability for flight speed, in contrast with these results and those at Struan.

Factors affecting temperament measurement

Previous experience

In the section on repeatability it was shown that lambs that were weighed frequently, starting before weaning had significantly lower CV than lambs that had been weighed less often. This difference continued or increased for at least 2 months after the more experienced lambs and less experienced had been run in the same mob.

Other factors

A limited number of factors have been tested by changing the conditions during weighing for some groups of sheep (Table 19). These were tested on the 2005 drop ewes, comparing the CV during this variation study with the CV recorded on four previous occasions

Table 19. Factors affecting CV in 2005 drop ewes

Factor	Coeff	SE	n	p%
Person in Front	-0.32	0.13	10	1.7
Different person weighing	-0.20	0.10	20	4.4
Sheep in Front	0.10	0.09	27	26
Sheep moving alongside	-0.20	0.17	6	23
Normal			74	

The operator who normally moves the sheep up to the working area stood in front of the scale during weighing. This significantly reduced the CV of the sheep being weighed. Therefore we recommend that bystanders do not stand in the area in front of the sheep while measurements are taken.

There was also a reduction in CV when this operator (who normally moved the sheep) weighed the sheep instead of the usual operator.

Sheep standing in front of the crate (about 5m ahead) did not significantly affect the result.

Sheep running backwards and forwards in the area adjacent to the weighing crate (about 6m to the side) did not appear to affect the results.

Some of these tests may not have included enough sheep to show a significant affect.

In all these studies dogs have been kept out of the weighing area. Tests will be needed on the effect of dogs in the vicinity, dogs barking in the distance etc.

Comparison of CV of weight with visual scores

The visual score methods developed by the South Australian group were compared with the CV of weight method. This will be reported in more detail later. The main results were that the visual behaviour scores correlated closely with CV, whereas flight time was not related to the other measures and that there were very marked differences between sire groups.

Conclusions

Several methods of measuring sheep temperament are described here. The repeatability appears to be sufficiently high so that for most purposes only one or two readings would be required to obtain a useful result. Temperament scores are related to some production traits with consistent benefits for calmer sheep. These traits include reproduction, growth and handling ability.

The visual behaviour score, settling time and 'jumpiness', measured on the scal, are all closely related, whereas flight speed is not related to these and some of the measures obtained using the dual readout system may also be different.

The visual scores are simple to use and do not require special equipment, but may depend on operator training or experience.

Flight time requires special equipment and there are a limited number of devices able to measure it, but it does not require the extended weighing time required by most of the other methods and it measures a different aspect of behaviour not correlated with the other measures studied here. Setting up the beams can be fiddly, but there are alternative methods of measuring flight time or flight speed, using simple technology such as radar guns, laser measuring devices etc.

The jumpiness score requires an electronic scale, but it has been assumed that most of these scoring systems would be used during normal weighing procedures. Although the scale readout must be connected directly to a computer, the calculation could be done by programs inside the readout unit. The current method includes magnetic switches on the gates and these require a special keyboard to transfer the signals to the computer. However, it is expected that the method would work without the switches, since weighing systems are able to determine when the sheep is ready to be weighed without such devices.

The double-readout method uses two readouts, one for each load cell. However, it is possible that a single readout could separate the signals from each load cell if this was necessary. The method appears to provide information on two additional aspects of behaviour.

Therefore we expect that an inexpensive, simple system of measuring temperament can be developed when the relationships between these measurements and production traits are more clearly established. This may require testing of more Information Nucleus Flocks using a selection of the current methods

Further work.

Pregnancy scanning and lamb survival

There are a variety of causes of lamb losses so measurement of effects of ewe temperament on mothering ability and lamb survival may require tests on a larger range of flocks combined with pregnancy scanning and udder checks at marking, or more detailed observation.

Sheep familiarity with the weighing system

The older ewes had been weighed many times in the past, but their familiarity with the weighing unit did not reduce the repeatability of temperament. Therefore additional studies will be needed on sheep of various ages that have not been weighed frequently so the experience is novel to the sheep being tested.

Other behavioural traits

Interest in sheep temperament was initially a result of the possible relationship with lamb mortality. However, some sheep are difficult to handle because they leave the mob when driven, jump fences, are difficult to handle in the race or during shearing. Additional behavioural observations could be made to determine whether there are other benefits or any disadvantages in selection for a quiet temperament. Comparisons may be needed with established methods such as the arena test.

Other production traits

In addition to lamb survival discussed above, temperament may be related to body weight, condition score, feed efficiency and meat quality. These effects should be examined in young growing sheep. The Information Nucleus Flocks should provide the information required.

Factors affecting the measurement

A limited number of factors affecting the measurement of temperament have been tested. It is necessary to determine external factors that may influence the measurement of sheep behaviour so that interfering events are avoided during weighing. This may include changes in the operator, other bystanders, the presence of other sheep, movement of other sheep in the shed, dogs barking, other noises, weather, time of day, length of time or distance away from paddock, etc.

Optimum age to test the sheep

The repeatability of the CV (and probably other measures) increases with age or possibly with experience of the sheep being weighed. Therefore repeated measurements and handling do not interfere with the ability to measure temperament. However, if selection of sheep on the basis of temperament is required it is necessary to avoid learned behaviour effects. Therefore it will be essential to determine whether the greater consistency with age is also linked to greater heritability for measurements in older sheep or whether heritability decreases with age.

Standard protocol

The current system includes dual readout units, timing beams, magnetic switches on the gates and an operator recording observations on the sheep.

The aim of this project is to develop a system that is simple, reliable, inexpensive, and provides results related to other performance characteristics. It is planned to have a system that will require only a weigh unit, tag reader and a standard portable computer for testing on other INF flocks. For commercial use the program is expected to be built into the weigh unit.

References

Beausoleil, NJ, Blache, D, Stafford, KJ, Mellor, DJ and Noble, ADL (2008) Exploring the basis of divergent selection for ‘temperament’ in domestic sheep. Applied Animal Behaviour **109**: 261-274

Blache, D. and Ferguson, D. (2005) Boost lamb survival – select for calm ewes. Sheep Updates, Department of Agriculture, WA.

Fell, L.R., Colditz, I.G., Walker, K.H. and Watson, D.L. (1999) Associations between temperament, performance and immune function in cattle entering a commercial feedlot. Aust. J. Expt. Ag. **39**: 795-802.

Forkman, B, Boissy, A, Meunier-Salaun, M-C, Canali, E and Jones, RB (2007) A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. Physiology and Behaviour, **92**: 340-374.

Gelez, H, Lindsay, DR, Blache, D, Martin, GB and Fabre-Nys, C (2003) Temperament and sexual experience affect female sexual behaviour in sheep. Applied Animal Behaviour **84**: 81-87.

Kilgour, RJ (1998) Arena behaviour is a possible selection criteria for lamb-rearing ability; it can be measured in young rams and ewes. Applied Animal Behaviour **57**: 81089.

Kilgour, RJ and Szantar-Coddington, MR (1995) Arena behaviour of ewes selected for superior mothering ability differs from that of unselected ewes. Animal Reproduction Science, **37**: 133-141.

Murphy, P.M., Purvis, I.W., Lindsay, D.R., Le Neindre, P., Orgeur, P. and Poindron, P. (1994) Measures of temperament are highly repeatable in merino sheep and some are related to maternal behaviour. Proc. Aust. Soc. Anim. Prod., **20**: 247-250.

Sart, S (2005) Temperament and milk quality in sheep and cattle. MSc Thesis University of Western Australia.

Starbuck, T.M., Hocking Edwards, J.E. and Hinch, G.N. (2006) Measurement of temperament/fearfulness of lambs without the use of specialist equipment. Proc. Aust. Soc. Anim. Prod. **26**: 43.