DELIVERABLE

DL.3.2.2: REPORT OF THE STUDY ON TEST INTER-OBSERVER REPEATABILITY AND FEASIBILITY OF INDICATORS OF CONSCIOUSNESS IN BROILER CHICKEN AFTER WATERBATH STUNNING AND THE IMPACT OF ELECTRICAL KEY PARAMETERS ON STUNNING EFFICIENCY.

Content

1. Introduction

Activity 3. Scientific and Technical Studies

Article 96 (d): Carrying out scientific and technical studies on the welfare of animals used for commercial or scientific purposes.

Sub-activity 3.2: Scientific and technical studies to validate indicators and methods

Objectives:

1. To help the development of indicators and methods for welfare assessment concerning the four priority areas.
2. To address some negative welfare aspects identified, in order to provide technical solutions to improve animal welfare.
3. To answer some queries of the CAs.

DELIVERABLE: DL.3.2.2. – REPORT OF THE STUDY ON TEST INTER-OBSERVER REPEATABILITY AND FEASIBILITY OF INDICATORS OF CONSCIOUSNESS IN BROILER CHICKEN AFTER WATERBATH STUNNING AND THE IMPACT OF ELECTRICAL KEY PARAMETERS ON STUNNING EFFICIENCY.

This deliverable is part of the sub-activity 3.2. "Scientific and technical studies". This study is aimed at investigating the pertinence of different animal-based indicators (ABIs) for the assessment of broiler's state of consciousness after waterbath stunning (WBS) before and during bleeding. The inter-observer repeatability of these indicators was studied in order to identify refined and validated ABIs with good level of repeatability that can be used for the assessment of the state of consciousness in commercial slaughterhouses. Moreover, the correlation among the outcomes of the ABIs and the effectiveness of stunning according to different combinations of waterbath electrical key parameters (frequency and current) used in different commercial slaughterhouses will be assessed.

2. Methods

2.1. Selection of slaughterhouses and animals

Six commercial broiler chicken slaughterhouses (SH) equipped with WBS were selected in France and Spain. Selection of the SH was carried together with the official veterinary services and reflect a certain diversity in terms of size of the plant, electrical key parameters, chicken strains, and line speed. Each slaughterhouse was assigned to a number (SH from 1 to 6).

2.2. Description of the slaughterhouses and waterbath stunning systems

Substantial variation of age, design and construction of the SH were observed, and the main characteristics are shown in Table 1.

Table 1. Main characteristics of the six slaughterhouses included in the study.

	Slaughterhouse					
	1	2	3	4	5	6
Location	France	France	France	Spain	Spain	Spain
Waterbath length (m)	0.9	6.0	3.0	3.7	3.8	3.3
Birds in the waterbath (n)	3	39	11	16	18	12-14
Exposure time (s)	30	14	9	13	15	11
Line speed (birds/h)	200	9,500	10,500	6,000	6,100	6,000
Mean time from the exit of the waterbath until bleeding (s)	1	2	3	9	11	6
Bleeding method*	M	A	A	MA	MA	M

*Bleeding method: M (manually); A (mechanically); MA (combination of first mechanically and afterwards manually)

In all SH, broilers were individually hung upside down by the legs on the moving shackles of the slaughter line and stunned by immersion of the head in the electrified waterbath. None of them had adjustable shackles to different weight and size of the broiler legs. The height of the waterbath was adjusted according to the size of the birds to facilitate all birds an immersion up to the base of their wings. Line speed was not measured *in situ*. The values are those reported by the food business operators and the official veterinary service. A digital control panel monitored the electrical parameters applied (*i.e.*, actual total current amount passing through the waterbath, the voltage and the frequency) in all SH. The automatically recorded electrical parameters were obtained from the slaughterhouse but were not measured and verified. The average values of current per animal was calculated by dividing the total current amount passing through the waterbath by the number of birds simultaneously in the water. The electrical waveform was sine alternating current in all the SHs. The bleeding procedure differed among SH. Hence, two of them did manual bleeding by cutting the carotids through an oropharynx incision (SH-1 and 6); two of them did it mechanically (automatic neck cutter; SH-2 and 3) and two of them did a combination of mechanical and manual (SH-4 and 5) as the automatic neck cutter only sectioned one of the carotids and operators sectioned manually the second carotid afterwards. Slaughter line speed ranged from 200 to 10,500 birds/h.

2.3. Assessment of the consciousness

2.3.1 Observers

The assessment of the stunning effectiveness was carried out by three trained observers. Each observer (Obs) was named as letter (A to C). An additional person randomly selected and identified the birds to be assessed by pointing at the bird with a laser pointer to prevent mistakes at evaluating all three observers the same selected bird. The stunning effectiveness was assessed in two different places of the slaughter line; 1) at the exit of the waterbath before bleeding and 2) during bleeding at approximately 10 s after severing the carotids, (Figure 1) in a representative sample of birds in each batch. The three observers assessed the bird and scored the ABIs between 3 and 6 s, depending on the slaughterhouse design and visibility. Observers assessed the ABIs independently and did not discuss or disclose their assessments during the evaluation.

2.3.2 Sample assessment

All the batches of broiler chickens slaughtered during the presence of the observers in the plant were evaluated. On each batch, samples of 50-100 birds were assessed before and during bleeding. This cycle was repeated until the whole batch was slaughtered, in order to reach the biggest sample size possible.

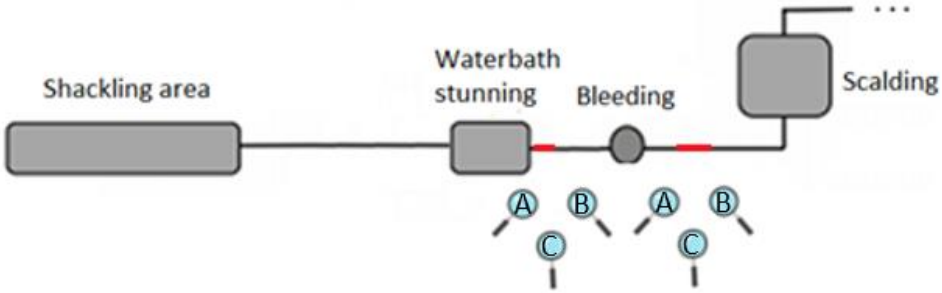

Figure 1. Position of the observers (A to C) during the assessment of ABIs of the effectiveness of waterbath stunning in broilers. The position of the lens is the position of the observers (*i.e.*, before and during bleeding) and the red segments are the observation area.

A summary of the electrical parameters used per batch and per SH along with the characteristics of the animals in the batch and the number of assessed birds is shown in Table 2.

Table 2. Number of batches, of slaughtered broilers, type of strain, average body weight and age of broilers per batch for each slaughterhouse. The number of broilers assessed before and during bleeding, the average electrical parameters ± standard deviation of the waterbath are also reported.

		Characteristics of the birds				No. Birds assessed		Electrical parameters		
SH	Batch	Strain	No. Birds	BW, kg	Age, d	BB	DB	Current, mA/bird	Frequency, Hz	Voltage, V
1	1	SG	95	2.8	134	55	50	102±15	60	80
	2	SG	57	2.7	107	46	226	105±26	60	80
	3	SG	311	2.1	104	161	161	112±29	60	80
2	1	FG	37,047	2.100	38	200	65	228±33	793	211
	2	FG	23,647	2.260	37	200	200	273±53	792	208
	3	FG	18,465	2.140	38	150	39	226±64	607	209
	4	FG	11,693	1.835	37	200	239	177±62	568	194
	5	SG	7,280	1.860	37	200	200	234±46	607	209
	6	SG	7,280	1.860	37	50	50	352±33	1507	279
3	1	FG	8,000	1.860	34	200	128	157±7	110	105
	2	FG	14,300	1.920	34	200	401	162±8	110	95
4	1	FG	3,240	2.843	44	104	0	309±43	347±1	68±13
	2	FG	3,888	2.837	44	100	124	276±71	348±1	60±22
	3	FG	3.888	2.924	44	50	150	255±59	349±1	65±17
5	1	FG	1,458	2.092	35	41	0	224±0	352±1	195±12
	2	FG	5,832	2.047	35	200	234	223±3	352±1	198±12
	3	FG	2,916	1.763	35	100	131	223±0	352±1	206±13
	4	FG	2,916	1.763	35	100	150	224±0	352±1	215±14
6	1	FG	2,700	3.560	56	100	0	105 to 141	80	33
	2	FG	720	4.020	60	0	38	105 to 111	80	32
	3	FG	4,350	3.570	55	0	200	105 to 143	80	32
	4	FG	6,300	2.830	43	100	195	108 to 153	80	32
	5	FG	5,040	3,080	50	100	100	106 to 160	80	33

N° Birds: number of birds in the batch; SG: slow growing; FG: fast growing; BW: body weight; SH: slaughterhouse; BB: before bleeding; DB: during bleeding.

2.3.3. Indicators for the assessment

The ABIs for the assessment of the state of consciousness before and during bleeding were selected based on those proposed by the EFSA (EFSA, 2013). The selected ABIs before bleeding were tonic seizure, breathing, spontaneous blinking and vocalisations, while the selected ones during bleeding were wing flapping, breathing, spontaneous swallowing and head shaking. The description and the outcome of consciousness and unconsciousness of these ABIs is summarized in Table 3.

Table 3. Animal-based indicators (ABI) assessed and descriptions of the outcomes of unconsciousness and consciousness in broilers stunned by waterbath in two different stages: before and during bleeding. Adapted from EFSA (2013).

Stage	ABI	Outcome of unconsciousness	Outcome of consciousness
Before bleeding	Tonic seizure	Bird shows general loss of muscle tone and a completely relaxed body and flaccid body, with no neck tension.	Bird shows arched and stiff neck (*i.e.*, necks appear parallel to the ground) and wings held tightly close to the body.
	Breathing	Absence of movements of the beak or abdominal muscles around the cloaca associated to cessation of breathing.	Presence of either a minimum of two movements of the beak or abdominal muscles around the cloaca associated to breathing.
	Spontaneous blinking	Bird does not open/close eyelid on its own (fast or slow) without stimulation.	Bird opens/closes eyelid on its own (fast or slow) without stimulation.
	Vocalisations	Absence of single or repeated short and loud shrieking (screaming) at high frequencies.	Single or repeated shrieking (screaming).
During bleeding	Wing flapping	Absence of flapping with both wings.	Flapping with both wings and should not be confused with rapid trembling of the entire body of the bird.
	Breathing	Absence of movements of the beak or abdominal muscles around the cloaca associated to cessation of breathing.	Either a minimum of two movements of the beak or abdominal muscles around the cloaca associated to breathing.
	Spontaneous swallowing	Absence of deglutition reflex.	Deglutition reflex triggered by water from the stunner or blood from the neck-cutting wound entering the mouth during bleeding.
	Head shaking	Bird does not shake its head from side to side.	Bird shakes its head from side to side to get rid of blood or water entering the nostrils.

The three trained observers agreed beforehand on the definition of the indicators, the methodology of assessment and the scoring to standardize the protocol when assessing the birds with these indicators.

Before the assessment of the birds, the three assessors were placed where there was the best possible visibility towards the shackled birds from a ventral position. However, due to divergence in the design and construction of the SH, sometimes the birds were assessed from a dorsal position instead of ventral (SH-5 and SH-6 at the exit of the waterbath and in SH-3 during bleeding) and thus, impairing the assessment of breathing by direct observation of cloacal rhythmic movements. Data were recorded as binomial as 0 if the outcome of unconsciousness was observed and 1 when an outcome of consciousness was observed. The presence of at least one outcome of consciousness may indicate the bird is conscious or regaining

consciousness after WBS and therefore of an ineffective stunning or a long stun-to-stick interval (*i.e.*, time from stunning to the start of bleeding).

2.4. Statistical analysis

Data pre-processing, statistical analyses and plots were performed using R software v.4.1.0. (R Core Team, 2021). First, birds that were not assessed by all three observers were filtered out to ensure that all observations were directly comparable. For all the statistical analyses, significance was declared at $P < 0.05$.

2.4.1. Inter-observer repeatability of ABIs

The overall level of agreement between observers for each ABI were determined and expressed by the crude proportion of agreement (PoA) and the Fleiss' kappa using the "irr" package of R software (Gamer *et al*., 2019). The PoA can be misleading as it does not take into account the scores that the raters assign due to chance. Fleiss's Kappa overcomes this issue as it provides an inter-observer agreement measure between two or more observers when the variable assessed is on binomial or categorical scale. It expresses the degree to which the observed proportion of agreement among observers exceeds what would be expected if all observers made their ratings completely randomly. Kappa can range from −1 to +1, where 0 indicates the amount of agreement that can be expected from random chance, and 1 represents perfect agreement between the observers (McHugh, 2012). Kappa is a standardized value and thus is interpreted the same across multiple studies. Thus, according to Fleiss *et al*. (2003) Fleiss' kappa can be classified as "excellent" agreement beyond chance if values are greater than 0.75; "fair to good" agreement beyond chance if values between 0.40 and 0.75 and "poor" agreement beyond chance if the values are below 0.40. However, when there is an insufficient scoring variation in the evaluated indicator (*i.e*., low prevalence of indicators of consciousness), although high agreement between observers, kappa appears close to 0.

2.4.2. Correlation among ABIs

As data did not follow a normal distribution, Spearman's rank correlation was performed to measure the association between the observed ABIs. Correlation results were displayed as heat map. Proportions among combinations of ABIs were performed as Venn diagram considering all broilers assessed in the present study using the "eulerr" package (Larsson, 2020).

2.4.3. Relationship between electrical parameters and stunning efficiency

Stunning inefficiency of each batch and SH was evaluated by showing the percentage of birds with at least one outcome of consciousness in any of the stages of the assessment: before and during bleeding. Chi-squared test was used to determine if there were statistical differences among observers between the expected and the observed frequencies of every outcome of the indicators evaluated. If one observer differed statistically from the others at evaluating the ABIs, the mean of the proportion of the two closest evaluations or the in between value when scoring were not consistent among them were reported. Moreover, confidence interval was computed taking into consideration the number of birds evaluated in each combination of electrical parameters used.

Moreover, a generalized linear mixed-effects model using the "lme4" package (Bates *et al*., 2015) was used to compare the effectiveness of stunning among the different combination of electrical key parameters. In

this statistical model, the slaughterhouse was included as fixed effect and the observer and the batch as random effects. Pairwise comparisons were Tukey adjusted using the package "emmeans" (Lenth *et al.*, 2021).

3. Results

The ABIs were assessed on a total of 2,685 broilers before bleeding and 3,154 during bleeding from 6 different SH across France and Spain by 3 observers. However, not all of them were assessed by the 3 observers. Those not assessed by all three observers were filtered out. Thus, 2,608 broilers remained in the dataset before bleeding and 3,105 during bleeding. The number of the birds assessed per SH as well as the number and percentage of birds assessed by the three observers is shown in Table 4.

Table 4. Number of assessed animals and number and percentage of birds that were able to be assessed by the three observers according to the slaughterhouse (before and during bleeding).

	Slaughterhouse						
Birds assessed	**1**	**2**	**3**	**4**	**5**	**6**	**All**
Before bleeding							
Total number of birds assessed	189	984	400	254	441	417	2,685
Number of birds assessed by the three observers	114	984	400	254	441	415	2,608
Birds assessed by the three observers, %	60.3	100.0	100.0	100.0	100.0	99.5	97.1
During bleeding							
Total number of birds assessed	209	793	529	374	516	733	3,154
Number of birds assessed by the three observers	195	778	527	374	515	716	3,105
Birds assessed by the three observers, %	93.3	98.1	99.6	100.0	99.8	97.7	98.4

3.1. Inter-observer repeatability of the ABIs

3.1.1. Before bleeding

After WBS and before bleeding, four ABIs of the state consciousness were assessed: tonic seizure breathing, spontaneous blinking and vocalisation. The overall level of agreement between the three observers for these ABIs according to the SH is shown in Table 5. Since SH-2 modified the electrical parameters in each batch of broilers, the overall level of agreement for this specific SH is also shown in Table 6. On the other hand, the prevalence of birds showing outcomes of consciousness, by observer and SH is shown in Table 7.

Table 5. Inter-observer proportion of agreement (PoA), 95% confidence interval (CI), Fleiss' kappa coefficient and interpretation, standard error (SE) of the animal-based indicators for the state of consciousness before bleeding in broilers according to the slaughterhouse assessed.

	Slaughterhouse						
Item	**1**	**2**	**3**	**4**	**5**	**6**	**All**
Tonic seizure							
PoA, % (95% CI, %)	91.2 (84.5 - 95.7)	97.4 (96.2 - 98.3)	99.5 (98.2 - 99.9)	99.2 (97.2 - 99.9)	96.6 (94.5 - 98.1)	61.0 (56.1 - 65.7)	91.7 (90.6 – 92.7)
Fleiss' Kappa (SE)	0.06 (0.05)	0.65 (0.02)	0.60 (0.03)	0.00 (0.04)	-0.01 (0.03)	0.47 (0.03)	0.64 (0.01)
Kappa interpretation	Poor	Fair to good	Fair to good	Poor	Poor	Fair to good	Fair to good
P-value	0.131	<0.0001	<0.0001	0.529	0.700	<0.0001	<0.0001
Breathing							
PoA, % (95% CI, %)	93.0 (86.6 - 96.9)	99.1 (98.3 - 99.6)	99.8 (98.6 - 100)	100 (98.8 - 100)	100 (99.3 - 100)	97.4 (95.3 - 98,7)	98.9 (98.4 - 99.3)
Fleiss' Kappa (SE)	0.77 (0.05)	0.25 (0.02)	0.00 (0.03)	*	*	0.14 (0.03)	0.58 (0.01)
Kappa interpretation	Excellent	Poor	Poor	*	*	Poor	Fair to good
P-value	<0.0001	<0.0001	0.512	*	*	<0.0001	<0.0001
Spontaneous blinking							
PoA, % (95% CI, %)	98.3 (93.8 - 99.8)	99.7 (99.1 - 99.9)	100 (99.3 - 100)	100 (98.8 - 100)	99.8 (98.7 - 100)	100 (99.3 - 100)	99.8 (99.5 - 99,9)
Fleiss' Kappa (SE)	0.00 (0.05)	0.00 (0.02)	*	*	0.05 (0.03)	*	0.14 (0.01)
Kappa interpretation	Poor	Poor	*	*	Poor	*	Poor
P-value	0.540	0.522	*	*	<0.0001	*	<0.0001
Vocalisation							
PoA, % (95% CI, %)	100 (97.4 - 100)	100 (99.7 - 100)	99.8 (98.6 - 100)	100 (98.8 - 100)	100 (99.3 - 100)	100 (99.3 - 100)	100 (99.8 - 100)
Fleiss' Kappa (SE)	*	*	*	*	*	*	*
Kappa interpretation	*	*	*	*	*	*	*
P-value	*	*	*	*	*	*	*

* Insufficient scoring variation to calculate kappa coefficients (all indicator scores were 0). Kappa interpretation: ≥ 0.75 'excellent', 0.40–0.74 'fair to good', and < 0.40 'poor' agreement (Fleiss *et al.*, 2003).

Table 6. Inter-observer proportion of agreement (PoA), 95% confidence interval (CI), Fleiss' kappa coefficient and interpretation, standard error (SE) of the animal-based indicators for the state of consciousness before bleeding in broilers in slaughterhouse 2 according to the batch assessed.

	Batch					
Item	**1**	**2**	**3**	**4**	**5**	**6**
Tonic seizure						
PoA, % (95% CI, %)	93.7 (89.2 - 96.7)	96.5 (92.9 - 98.6)	100 (98.0 – 100)	98.0 (95.0 - 99.5)	99.5 (97.2 – 100)	95.9 (86.0 – 99.5)
Fleiss' Kappa (SE)	0.46 (0.04)	0.80 (0.04)	*	0.19 (0.04)	0.00 (0.04)	0.74 (0.04)
Kappa interpretation	Fair to good	Excellent	*	Poor	Poor	Fair to good
P-value	<0.0001	<0.0001	*	<0.0001	0.516	<0.0001
Breathing						
PoA, % (95% CI, %)	97.9 (94.7 - 99.4)	98.5 (95.6 - 99.7)	98.7 (95.2 - 99-8)	100.0 (98.5 – 100)	100 (98.5 – 100)	100 (94.1 – 100)
Fleiss' Kappa (SE)	0.19 (0.04)	-0.01 (0.04)	0.50 (0.05)	*	*	*
Kappa interpretation	Poor	Poor	Fair to good	*	*	*
P-value	<0.0001	0.549	<0.001	*	*	*
Spontaneous blinking						
PoA, % (95% CI, %)	100 (98.4 – 100)	99.5 (97.2 – 100)	99.3 (96.3 – 100)	100 (98.5 – 100)	100 (98.5 – 100)	98.0 (89.2 – 99.9)
Fleiss' Kappa (SE)	*	-0.00 (0.04)	-0.00 (0.05)	*	*	0.00 (0.08)
Kappa interpretation	*	Poor	Poor	*	*	Poor
P-value	*	0.516	0.519	*	*	0.533
Vocalisation						
PoA, % (95% CI, %)	100 (98.4 – 100)	100 (98.5 – 100)	100 (98.0 – 100)	100 (98.5 – 100)	100 (98.5 – 100)	100 (94.1 – 100)
Fleiss' Kappa (SE)	*	*	*	*	*	*
Kappa interpretation	*	*	*	*	*	*
P-value	*	*	*	*	*	*

* Insufficient scoring variation to calculate kappa coefficients (all indicator scores were 0). Kappa interpretation: ≥ 0.75 'excellent', 0.40–0.74 'fair to good', and < 0.40 'poor' agreement (Fleiss *et al.*, 2003).

Table 7. Percentage of the outcomes of the animal-based indicators for the state of consciousness in broilers after waterbath stunning but before bleeding (TS: tonic seizure; BR: breathing; SB: spontaneous blinking; VC: vocalisation) according to the observer (Obs: A to C) and slaughterhouses (SH) assessed.

		Absence of TS, %					Presence of BR, %					Presence of SB, %					Presence of VC, %				
SH	Birds, n	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value
1	114	2.6	6.1	0.9	3.5	0.072	12.3	11.4	11.4	11.4	0.972	0.9	0.9	0.0	0.9	0.605	0.0	0.0	0.0	0.0	-
2	984	3.0	2.6	2.1	2.6	0.443	0.7	0.3	0.2	0.4	0.173	0.2	0.0	0.0	0.0	0.368	0.0	0.0	0.0	0.0	-
3	400	0.8	0.3	0.3	0.5	0.448	0.0	0.3	0.0	0.0	0.368	0.0	0.0	0.0	0.0	-	0.3	0.0	0.0	0.0	0.366
4	254	0.0[b]	0.8[a]	0.0[b]	0.0	0.049	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-
5	441	3.4[a]	0.0[b]	0.0[b]	0.0	<0.001	0.2	0.2	0.2	0.2	1.000	0.2	0.0	0.2	0.2	0.606	0.2	0.2	0.2	0.2	1.000
6	415	36.4[b]	36.6[b]	59.8[a]	36.6	<0.001	0.7	1.4	1.0	1.0	0.580	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-
All	2608	7.7[b]	7.2[b]	10.4[a]	7.5	<0.001	1.0	0.9	0.8	0.9	0.736	0.2	0.0	0.1	0.1	0.368	0.1	0.0	0.0	<0.1	0.779

a–b = Values with different superscripts within the same raw differ among observers by chance ($P < 0.05$).

Table 8. Percentage of the outcomes of the animal-based indicators for the state of consciousness in broilers after waterbath stunning but before bleeding (TS: tonic seizure; BR: breathing; SB: spontaneous blinking; VC: vocalisation) according to the observer (Obs: A to C) in the slaughterhouses B for every batch assessed.

		Absence of TS, %					Presence of BR, %					Presence of SB, %					Presence of VC, %				
SH	Birds, n	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value
1	189	6.9[a]	3.7[ab]	1.6[b]	3.7	0.032	1.6	0.5	0.5	1.1	0.446	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-
2	198	6.1	5.6	7.1	6.1	0.817	1.0	0.0	0.5	0.5	0.366	0.5	0.0	0.0	0.0	0.606	0.0	0.0	0.0	0.0	-
3	149	0.7	0.7	0.7	0.7	1.000	1.3	1.3	0.0	0.7	0.134	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-
4	200	0.5	2.0	0.0	1.0	0.073	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-
5	199	0.0	0.5	0.0	0.0	0.367	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-
6	49	6.1	4.1	6.1	6.1	0.876	0.0	0.0	0.0	0.0	-	2.0	0.0	0.0	0.0	0.365	0.0	0.0	0.0	0.0	-

a–b = Values with different superscripts within the same raw differ among observers by chance ($P < 0.05$)

3.1.1.1. Tonic seizure

Birds with absence of tonic seizure at the exit of the waterbath was observed in SH 1, 2, 3 and 6. However, there was divergence in its prevalence according to the SH assessed (Table 7). While SH-3 did not exceed the 0.5% in average among observers, SH-6 had the highest prevalence in the sample with an average of 36.6%. In any case, the PoA was above 91% in all the SH except for the SH-6 where it was lower (61%) due to divergence in scoring birds among observers. Actually, observer C scored 1.4 times more birds with absence of tonic seizure compared with the other observers ($P < 0.0001$). Moreover, the Fleiss' kappa (κ) strongly varied among SHs: SH-1, SH-4 and SH-5 being close to 0 and interpreted as "poor agreement" and SH-2, SH-3 and SH-6 being between 0.47 and 0.65 and interpreted as "fair to good" (Table 5). Those had a κ close to 0 reflects an insufficient scoring variation linked to low prevalence of birds showing absence of tonic seizure.

Considering the data from the total of birds assessed in the present study (n=2,608), the 7.5% of birds showed absence of tonic seizure (Table 7) and the PoA among observers was high (91.7%) and the Fleiss' kappa coefficient was statistically significant ($\kappa = 0.64$; $P < 0.0001$; Table 5).

3.1.1.2. Breathing

Birds with presence of breathing was observed in SH-1, SH-2 (batches 1, 2 and 3) and SH-6 (Table 7 and 8). The highest prevalence of breathing in a sample was found in SH-1 with an average of 11.4% (Table 7). In any case, the PoA was above 93% in all SHs (Table 5) and there was no divergence on rating among observers ($P > 0.05$) in any SH nor batch assessed (Table 7 and 8). However, there was divergence of κ linked to the different degree of prevalence of breathing among SHs (Table 5 and 6)

Taking all birds from the SHs assessed into consideration, presence of breathing was observed in the 0.9% of birds (Table 7), the PoA among observers was high (98.9%) and the Fleiss' kappa coefficient was statistically significant ($\kappa = 0.58$; $P < 0.0001$; Table 5).

3.1.1.3. Spontaneous blinking

Birds showing spontaneous blinking was observed in SH-1 and SH-5. However, the higher prevalence in a sample was found at SH-1 with an average of 0.9% of the broilers (Table 7). In any case, the PoA was above 98.0% (Table 5 and 6) and there was no divergence on rating among observers ($P > 0.05$) in any SH assessed (Table 7). Moreover, there were no divergence of κ and was usually close to 0 showing that the prevalence of spontaneous blinking is low considering all the birds assessed.

It should be highlighted that spontaneous blinking was observed in 0.1% of the total birds assessed (Table 7) with a high PoA among observers (more than 99.8%) but with low Fleiss' kappa coefficient although statistically significant ($\kappa = 0.14$; $P < 0.0001$; Table 5).

3.1.1.4. Vocalisation

Vocalisation was heard only in SH-5 with an average of 0.2% of the broilers (Table 7). In any case, among all ABIs assessed before bleeding, vocalisation was the one with the highest PoA (above 99.8%) and there was no divergence on rating among observers ($P > 0.05$) in all SH assessed.

Taking all birds assessed in this study into consideration, the PoA among observers was 100% but the Fleiss' kappa coefficient was not computed due to insufficient scoring variation. Detection of vocalisation was extremely low (0.04%; Table 7).

3.1.2. During bleeding

Four ABIs were evaluated during bleeding: wing flapping, breathing, spontaneous swallowing and head shaking. The overall level of agreement between the three observers for these ABIs according to the SH is shown in Table 9. Since SH-2 modified the electrical parameters in each batch of broilers, the overall level of agreement for this specific SH according to the batch is shown in Table 10. On the other hand, the prevalence of birds showing indicators of consciousness according to the ABI per observer and SH assessed is shown in Table 11. As there was a considerable number of conscious birds at SH-2 and every batch of birds were stunned at different electrical key parameters, results of prevalence were split per batch at SH-2 and shown in Table 12.

3.1.2.1. Wing flapping

Birds with presence of wing flapping were observed in SH-2 (in batch 2, 4 and 5) and SH-5 and SH-6 (Table 11) with the highest prevalence being in SH-6 (5.0% broilers in average; Table 11). However, the prevalence of wing flapping strongly differed between these SHs and thus, so the κ and its interpretation. The PoA among observers was above 94% in all SHs and batches assessed (Table 9 and 10). Moreover, there was uniformity on rating among observers ($P > 0.05$) in all SHs except for the broilers assessed at SH-6 were one observer scored 2.3 times more wing flapping than the other observers ($P < 0.001$; Table 11).

Taking all birds from the SHs assessed into consideration, although the PoA among observers was high (98.2%) and the Fleiss' kappa coefficient was statistically significant ($\kappa = 0.66$; $P < 0.0001$; Table 9), the detection of wing flapping differed statistically among evaluators ($P < 0.01$) and thus, the prevalence was considered to be 1.6% as this was the in between value (Table 11).

3.1.2.2. Breathing

Birds with presence of breathing during bleeding were observed in birds of all SHs assessed but in SH-4 (Table 11 and 12). The highest prevalence occurred in batch 6 (38.8%), followed by SH-6 (36.9%) and some batches in SH-2 (batch 1, 33.9%; batch, 20.5%; batch 2, 15.7%; batch 4, 15.1%), in SH-1 (4.1%), SH-3 (1.7%) and SH-5 (0.4%).

Table 9. Inter-observer proportion of agreement (PoA), 95% confidence interval (CI), Fleiss' kappa coefficient and its interpretation and standard error (SE) of the animal-based indicators for the state of consciousness during bleeding according to the slaughterhouse assessed.

	Slaughterhouse						
Item	**1**	**2**	**3**	**4**	**5**	**6**	**All**
Wing flapping							
PoA, % (95% CI, %)	99.5 (97.2 - 100)	98.6 (97.5 - 99.3)	99.8 (99.0 - 100)	100 (99.2 - 100)	99.6 (98.6 - 99,9)	94.3 (92.3 - 95,9)	98.2 (97.7 - 98.6)
Fleiss' Kappa (SE)	0.00 (0.04)	0.26 (0.02)	0.00 (0.03)	*	0.00 (0.03)	0.66 (0.02)	0.66 (0.01)
Interpretation of kappa	Poor	Poor	Poor	*	Poor	Fair to good	Fair to good
P-value	0.517	<0.0001	0.510	*	<0.0001	<0.0001	<0.0001
Breathing							
PoA, % (95% CI, %)	93.0 (86.6 - 96.9)	85.2 (82.5 - 87.6)	97.7 (96.1 - 98.8)	100 (98.8 - 100)	100 (99.3 - 100)	97.4 (95,3 - 98,7)	88.2 (87.0 - 89.3)
Fleiss' Kappa (SE)	0.58 (0.04)	0.63 (0.02)	0.33 (0.03)	*	0.57 (0.03)	0.54 (0.02)	0.64 (0.01)
Interpretation of kappa	Fair to good	Fair to good	Poor	*	Fair to good	Fair to good	Fair to good
P-value	<0.0001	<0.0001	<0.0001	*	<0.0001	<0.0001	<0.0001
Spontaneous swallowing							
PoA, % (95% CI, %)	98.5 (95.6 - 99.7)	99.5 (93.8 - 96.9)	99.8 (98.9 -100)	100 (99.2 - 100)	100 (99.4 - 100)	98.5 (97.3 - 99.2)	98.4 (97.9 - 98.8)
Fleiss' Kappa (SE)	0.24 (0.04)	0.19 (0.02)	0.00 (0.03)	*	*	0.21 (0.02)	0.20 (0.01)
Interpretation of kappa	Poor	Poor	Poor	*	*	Poor	Poor
P-value	<0.0001	<0.0001	0.510	*	*	<0.0001	<0.0001
Head shaking							
PoA, % (95% CI, %)	99.5 (97.2 - 100)	92.9 (90.9 - 94.6)	99.4 (98.4 - 99.9)	100 (99.2 - 100)	100 (99.4 - 100)	92.6 (90.4 - 94.4)	96.4 (95.7 - 97.0)
Fleiss' Kappa	0.75 (0.04)	0.64 (0.02)	0.80 (0.03)	*	*	0.58 (0.02)	0.64 (0.01)
Interpretation of kappa	Excellent	Fair to good	Excellent	*	*	Fair to good	Fair to good
P-value	<0.0001	<0.0001	<0.0001	*	*	<0.0001	<0.0001

* Insufficient scoring variation to calculate kappa coefficients (all indicator scores were 0). Kappa interpretation: ≥ 0.75 'excellent', 0.40–0.74 'fair to good', and < 0.40 'poor' agreement (Fleiss *et al.*, 2003).

Table 10. Inter-observer proportion of agreement (PoA), 95% confidence interval (CI), Fleiss' kappa coefficient and its interpretation and standard error (SE) of the animal-based indicators for the state of consciousness during bleeding in waterbath stunned broilers at slaughterhouse 2 according to the batch assessed.

	Batch					
Item	**1**	**2**	**3**	**4**	**5**	**6**
Wing flapping						
PoA, % (95% CI, %)	98.4 (91.3 – 100)	96.5 (92.9 – 98.6)	100 (92.6 – 100)	98.7 (96.3 – 99.7)	100 (98.5 - 100)	100 (94.1 - 100)
Fleiss' Kappa (SE)	-0.01 (0.07)	-0.01 (0.04)	*	0.25 (0.04)	*	*
Kappa interpretation	Poor	Poor	*	Poor	*	*
P-value	0.529	0.614	*	<0.0001	*	*
Breathing						
PoA, % (95% CI, %)	62.9 (49.7 – 74.8)	84.9 (79.1 – 89.5)	82.1 (66.5 – 92.5)	84.1 (78.7 – 88.5)	96.0 (92.2 - 98.2)	79.6 (65.7 - 89.8)
Fleiss' Kappa (SE)	0.45 (0.07)	0.61 (0.04)	0.63 (0.09)	0.59 (0.04)	0.62 (0.04)	0.72 (0.08)
Kappa interpretation	Fair to good	Fair to good	Fair to good	Fair to good	Fair to good	Fair to good
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Spontaneous swallowing						
PoA, % (95% CI, %)	88.7 (78.1 – 95.3)	95.5 (91.6 – 97.9)	94.9 (82.7 – 99.4)	95.7 (92.2 – 97.9)	98.0 (94.9 - 99.5)	93.9 (83.1 - 98.7)
Fleiss' Kappa (SE)	0.09 (0.07)	0.08 (0.04)	0.32 (0.09)	0.32 (0.04)	0.19 (0.04)	-0.02 (0.08)
Kappa interpretation	Poor	Poor	Poor	Poor	Poor	Poor
P-value	0.121	0.020	<0.001	<0.0001	<0.0001	0.600
Head shaking						
PoA, % (95% CI, %)	77.4 (65.0 – 87.1)	93.4 (89.0 – 96.5)	89.7 (75.8 – 97.1)	92.7 (88.5 – 95.7)	98.0 (94.9 - 99.5)	93.9 (83.1 - 98.7)
Fleiss' Kappa (SE)	0.41 (0.07)	0.72 (0.04)	0.68 (0.09)	0.67 (0.04)	0.59 (0.04)	0.60 (0.08)
Kappa interpretation	Fair to good	Fair to good	Fair to good	Fair to good	Fair to good	Fair to good
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

* Insufficient scoring variation to calculate kappa coefficients (all indicator scores were 0). Kappa interpretation: ≥ 0.75 'excellent', 0.40–0.74 'fair to good', and < 0.40 'poor' agreement (Fleiss *et al.*, 2003).

Table 11. Percentage of the outcomes of the animal-based indicators for the state of consciousness in waterbath stunned broilers during bleeding (WF: wing flapping; BR: breathing; SB: spontaneous swallowing; HS: head shaking) according to the observer (Obs: A to C) and the slaughterhouses assessed (SH).

		Presence of WF, %					Presence of BR, %					Presence of SS, %					Presence of HS, %				
SH	Birds, n	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value
1	195	0.0	0.5	0.0	0.0	0.134	4.6	5.1	3.1	4.1	0.368	0.5	0.5	1.0	0.5	0.865	0.5	1.0	0.5	0.5	0.865
2	778	1.2	0.5	0.3	0.6	0.172	15.9	16.1	15.0	15.7	0.885	2.8	1.0	1.8	1.9	0.066	6.9	8.5	5.9	7.1	0.210
3	527	0.0	0.2	0.0	0.0	0.368	1.9^{a}	0.2^{b}	1.3^{a}	1.7	0.007	0.0	0.2	0.0	0.0	0.368	0.9	0.9	0.9	0.9	1.000
4	374	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-
5	515	1.4	1.2	1.4	1.4	0.957	0.8	0.4	0.2	0.4	0.195	0.0	0.0	0.0	0.0	-	0.0	0.0	0.0	0.0	-
6	716	8.1^{a}	5.3^{b}	4.6^{b}	5.0	<0.001	36.9^{b}	41.1^{a}	28.8^{c}	36.9	<0.001	0.4	0.6	1.0	0.7	0.193	6.7	6.6	5.6	6.3	0.437
All	3105	2.4^{a}	1.6^{ab}	1.4^{b}	1.6	0.006	13.2^{a}	13.9^{a}	10.9^{b}	13.6	<0.001	0.8	0.5	0.7	0.7	0.154	3.5	3.9	3.9	3.4	0.627

$^{a–c}$ = Values with different superscripts within the same raw differ among observers by chance ($P < 0.05$).

Table 12. Percentage of the outcomes of the animal-based indicators for the state of consciousness in waterbath stunned broilers during bleeding (WF: wing flapping; BR: breathing; SB: spontaneous swallowing; HC: head shaking) according to the observer (Obs: A to C) and the slaughterhouses B for every batch assessed.

		Presence of WF, %					Presence of BR, %					Presence of SS, %					Presence of HS, %				
Batch	Birds, n	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value	Obs-A	Obs-B	Obs-C	Mean	*P*-value
1	62	0.0	1.0	0.0	0.0	0.366	33.5	37.1	30.6	33.9	0.734	3.2	1.6	8.1	4.8	0.183	9.7^{b}	24.2^{a}	11.3^{b}	11.3	0.046
2	198	3.0^{a}	0.6	0.0^{b}	0.5	<0.001	15.7	15.2	15.7	15.7	0.987	1.5	1.0	2.5	1.5	0.491	9.1	10.1	6.1	8.6	0.321
3	39	0.0	0.0	0.0	0.0	-	15.4	23.1	23.1	20.5	0.131	5.1	0.0	2.6	2.6	0.358	12.8	15.4	7.7	12.8	0.567
4	232	0.9	0.4	0.4	0.4	0.363	18.1^{a}	14.7^{b}	12.9^{b}	15.1	0.028	4.3^{a}	0.9^{b}	1.3^{b}	1.3	0.021	8.6^{a}	7.8^{a}	7.8^{b}	8.2	0.925
5	198	0.5	0.5	0.5	0.5	1.000	4.0	3.0	4.0	3.5	0.828	1.0	1.5	0.0	1.0	0.244	1.5	1.5	2.0	1.5	0.903
6	49	0.0	0.0	0.0	0.0	-	30.6	46.9	40.8	38.8	0.248	6.1^{a}	0.0^{b}	0.0^{b}	0.0	0.047	4.1	8.2	4.1	6.1	0.589

$^{a–b}$ = Values with different superscripts within the same raw differ among observers by chance ($P < 0.05$).

The lower the prevalence of birds with presence of breathing, the higher the agreement among observers (as all observers agreed on the absence of ABIs) and the poorer the kappa interpretation. The higher the prevalence of birds with presence of breathing, the higher the likelihood of disagreement. The divergence in prevalence lead to divergence of κ linked to the different degree of prevalence of breathing among SHs and ranging from "poor" to "excellent" agreement according to Fleiss *et al.* (2003). On the other hand, divergence on rating breathing among observers on birds assessed were observed at SH-3, SH-6 and batch 4 from SH-2 ($P < 0.01$; Table 11 and 12).

Considering the data from all SH, the average PoA was 88.2% and the Fleiss' kappa coefficient was statistically significant ($\kappa = 0.64$; $P < 0.0001$; Table 9), the detection of breathing was observed in 13.6% of the assessed birds (Table 11).

3.1.2.3. Spontaneous swallowing

Birds showing spontaneous swallowing were observed in SH-1, SH-2 and SH-6. However, its prevalence was low (SH-1: 1.2%; SH-2: 1.5%; SH-6: 1.1%) compared with the presence of the outcomes of consciousness of other ABIs. For this reason, there was no divergence of κ and the level of agreement was classified as "poor" under every condition tested (Table 9 and Table 10) according to Fleiss et al. (2003). On the other hand, there was no divergence at scoring spontaneous swallowing among observers ($P > 0.05$; Table 11 and 12) except for the batch 4 from SH-2 ($P < 0.0001$).

When considering all the birds assessed, the PoA was very high (96.4%), but the Fleiss' kappa coefficient was low but statistically significant ($\kappa = 0.20$; $P < 0.0001$; Table 9) and the prevalence of spontaneous swallowing was 0.7% (Table 11).

3.1.2.4. Head shaking

Birds showing head shaking were observed in SH-1, SH-2, SH-3, SH-5 and SH-6 and the prevalence of head shaking varied among the SHs (SH-1: 0.5%; SH-2: 7.1%; SH-3: 0.9%; SH-5: 6.3% and SH-6: 3.4%). There was agreement at scoring head shaking among observers ($P > 0.05$; Table 11 and 12) except for the batch 4 from SH-2 ($P < 0.001$). Thus, the κ and the level of agreement was classified from "fair to good" to "excellent" according to the SH assessed (Table 9 and Table 10).

When taking all the birds assessed into consideration, the PoA among observers was very high (97.2%) and the Fleiss' kappa coefficient was statistically significant ($\kappa = 0.64$; $P < 0.0001$; Table 9), and its prevalence was found in the 3.8% of observed animals (Table 11).

3.2. Correlation among ABIs

3.2.1. Before bleeding

To elucidate the correlation among the outcomes of the ABIs assessed before bleeding, a contingency table was created. The proportions of birds showing outcomes of consciousness and their combinations observed at the same bird at this stage is shown as Venn diagram in Figure 2A. Absence of tonic seizure was the most frequent indicator followed by breathing. Spontaneous blinking, vocalisation and combinations between the

outcomes of consciousness of the four ABIs were almost non-existent at this stage. Heat map was not displayed in the report as no correlation was found among any ABI.

3.2.2. During bleeding

Contingency table showing the outcomes of the ABIs during bleeding was also created. The proportions of birds showing outcomes of consciousness and their combinations observed at individual level at this stage is shown as Venn diagram in Figure 2B. This diagram showed that presence of breathing was the most frequent outcome of consciousness observed followed by head shaking and spontaneous swallowing whereas the observation of wing flapping was rare. Additionally, when the prevalence of birds showing indicators of consciousness was high as occurred at SH-2 and SH-6, some birds showed breathing accompanied primordially by head shaking but rarely by spontaneous swallowing.

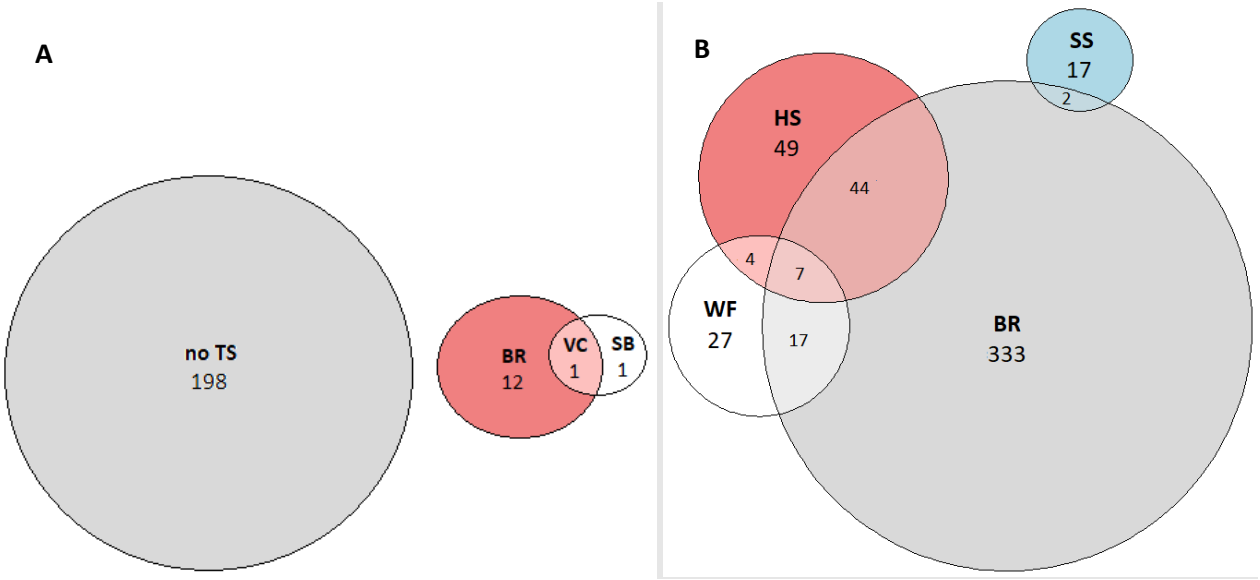

Figure 2. Venn diagram of the animal-based indicator of consciousness assessed in broilers A) before bleeding in waterbath stunned broilers and B) during bleeding. Indicators of consciousness were: no TS: absence of tonic seizure; BR: presence of breathing; SB: presence of spontaneous blinking; VC: presence of vocalisation; WF: presence of wing flapping; HS: presence of head shaking; SS: presence of spontaneous swallowing. Numbers specify the total amount of broilers showing each indicator or combinations of indicators from a total of 2,608 broilers assessed before bleeding and 3,105 during bleeding.

Unlike the before bleeding assessment, there was correlation among ABIs as shown as a heat map in Figure 3. All correlations were positive, but some were statistically significant such as the presence of breathing and wing flapping ($r = 0.71$; $P < 0.001$), breathing and head shaking ($r = 0.90$; $P < 0.001$) and head shaking and spontaneous swallowing ($r = 0.63$; $P = 0.02$). However, there was no correlation between wing flapping and spontaneous swallowing ($r = 0.22$; $P = 0.337$).

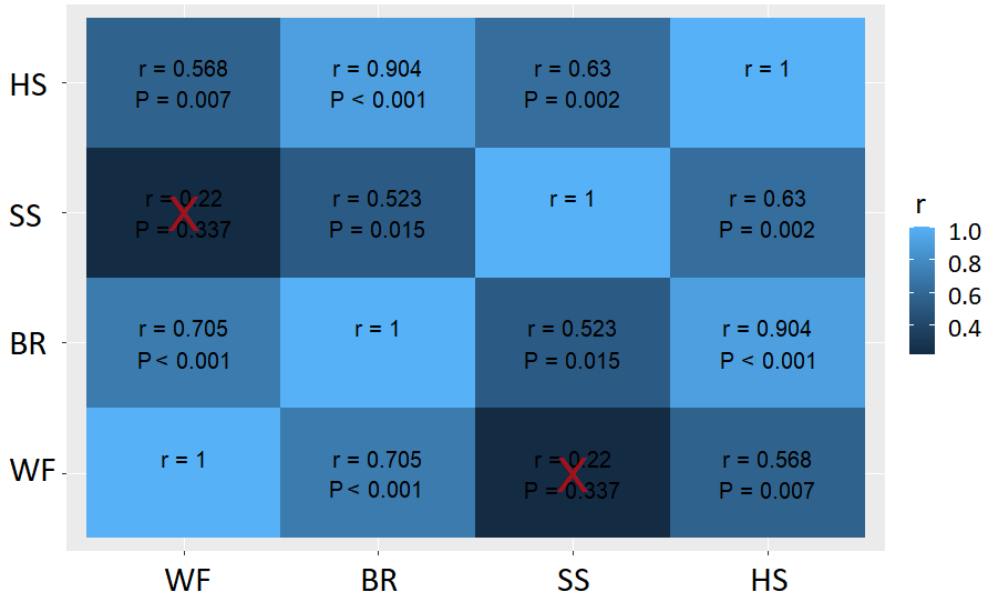

Figure 3. Heat map of correlations of the outcomes of the animal-based indicators for the state of consciousness during bleeding in waterbath stunned broilers (WF: wing flapping; BR: breathing; SS: spontaneous swallowing; HS: head shaking) found when assessing 19 batches from six different slaughterhouses. The values in the table are Spearman's rank correlation coefficients (r) and *P*-values. Red crosses indicate non-significant correlations ($P > 0.05$).

3.3. Relationship between electrical parameters and stunning effectiveness

Stunning effectiveness was analysed in relation to different combinations of electrical parameters applied to batches of different characteristics (*e.g.*, body weight, age, type of strain). To gain some insight on relationship between combination of electrical parameters and stunning effectiveness and maintenance of the state of unconsciousness in broilers, the prevalence of birds showing indicators of consciousness were compared. A summary of the different electrical parameters applied and the 95% confidence interval of birds showing at least one outcome of consciousness is summarized in Table 15. Statistical comparison of the odds to find broilers with outcomes of consciousness among SH is shown in Table 16.

The best combination of electrical parameters that resulted in effective stunning was found at SH-4 as any broiler did not show outcomes of consciousness. SH-4 was closely followed by SH-5 where the prevalence of conscious birds (showing at least one outcome of consciousness) ranged from 0 to 2% but the odds of finding a broiler with outcomes of consciousness significantly differed with those found at SH-4 ($P = 0.021$).

On the other hand, the prevalence of conscious birds in SH-1 (4.9 to 5.1%) was higher than in SH-4 (0.0%) and so the odds of effective stunning ($P < 0.001$) but equal to SH-5 ($P = 0.267$) and to batch 5 of SH-2 ($P = 1.000$). Far away from SH-4 and SH-5, the prevalence of conscious broilers was high and the odds of finding a conscious broiler was statistically equal for batch 1, 2, 3, 4, 6 from SH-2 and SH-6 ($P > 0.05$).

Table 15. Electrical parameters used in the waterbath, number of broilers assessed (n), prevalence and odds ratio of broilers showing at least one outcome of consciousness during bleeding according to the slaughterhouse (SH:1 to 6) and batch assessed.

			Electrical parameters in waterbath			Birds with outcomes of consciousness, %		
SH	**Batch**	**n**	**Current, mA/bird**	**Frequency, Hz**	**Voltage, V**	**Mean**	**CI 95%**	**Odds ratio**
1	1&2&3	195	106±23	60	80	4.8*	[1.7-7.6]§	0.101[b]
2	1	62	228±33	793	211	50.0**	[37.6-62.5]§§	-[a]
2	2	198	273±53	792	208	22.2**	[16.4-28.0]§§	0.697[a]
2	3	39	226±64	607	209	29.1**	[14.1-42.3]§§	1.240[a]
2	4	232	177±62	568	194	20.5**	[15.5-25.9]§§	0.472[a]
2	5	198	234±46	607	209	5.6**	[2.4-8.8]§§	0.122[b]
2	6	49	352±33	1507	279	41.5**	[27.1-54.6]§§	1.492[a]
3	1&2	527	200	128 to 401	160±7	2.0*	[0.9-3.3]§§	0.012[c]
4	1&2&3	374	280±57	358±1	64±19	0.0*	[0.0-0.0]§	0.000[d]
5	1&2&3&4	515	223±1	352±1	205±13	1.6*	[0.5-2.6]§	0.024[c]
6	1&2&3&4&5	716	105 to 141	80	32	41.2*	[37.6-44.8]§	1.610[a]

*: mean of all three observers per batch
**: mean of results of all batches of the slaughterhouse
§: 95% confidence interval for all batches and all observers
§§: 95% confidence interval for one batch and all observers
[a,d] Different letters in the same column indicate differences (P < 0.05) due to electrical parameters in waterbath

Table 16. Statistical comparison of the odds ratio of finding broilers showing at least one outcome of consciousness after waterbath stunning and during bleeding between slaughterhouses (SH: 1 to 6) and batch within SH-2 as differed on the electrical parameters applied in waterbath (see Table 15). Stressed in red are significant differences between flocks ($P < 0.05$).

SH-Batch	**1**	**2-1**	**2-2**	**2-3**	**2-4**	**2-5**	**2-6**	**3**	**4**	**5**	**6**
1	-										
2-1	<0.001	-									
2-2	<0.001	0.999	-								
2-3	<0.001	1.000	0.673	-							
2-4	<0.001	0.890	0.838	0.006	-						
2-5	1.000	0.008	0.002	<0.001	0.041	-					
2-6	<0.001	0.999	0.781	1.000	0.148	<0.001	-				
3	0.026	<0.001	<0.001	<0.001	<0.001	0.042	<0.001	-			
4	<0.001	<0.001	<0.001	<0.001	0.002	0.001	<0.001	0.023	-		
5	0.267	<0.001	<0.001	<0.001	<0.001	0.048	<0.001	1.000	0.021	-	
6	<0.001	0.995	0.001	0.984	<0.001	<0.001	1.000	<0.001	<0.001	<0.001	-

4. Discussion

Avoiding (or minimising) pain, fear and distress through a rapid induction of unconsciousness and death during slaughter is the main objective to be achieved from an animal protection point of view. An unconscious animal is insensitive to stimulations from the environment as the brain is no longer capable of dealing with sensory information (Terlouw *et al.*, 2016a). However, in WBS, although with the electric current parameters set up by the regulation, not all birds are successfully rendered unconscious and some of them may be ineffective stun or recover consciousness before death. For this reason, it is mandatory to check that all birds are unconscious from the exit of the waterbath until death occur through bleeding and exsanguination. EFSA (2013) provided a list of ABIs to evaluate the state of consciousness including the level of feasibility, sensitivity and specificity. However, the inter-observer repeatability of the outcome of this ABIs has not been assessed yet.

Thus, one of the aims of the study was to gain insight into the inter-observer repeatability of some feasible ABIs for the state of consciousness after WBS in broiler chickens. This is the first study that compares the assessment of three observers in 5,241 broilers from 19 batches of 6 different slaughterhouses and 11 different key electrical parameters applied in waterbath from two main broiler producer countries in the UE-27. In addition, it should be highlighted that not only observers were well trained, but they also agreed on the definition of the indicators before the assessment of the birds. In addition, SH were chosen to be of large variability in designs, key electrical parameters and line speeds to provide a good overview. The number of observers was intended to cause the minimum interference to the operators. Although there was a restriction on available space for the assessment, the observers stood next to each other assessing the same animals at the same span of time.

4.1. Inter-observer repeatability of ABIs

Data were analysed at individual broiler level and the combination of Fleiss' kappa and PoA was used to assess the inter-observer repeatability of the outcomes of some ABIs for the state of consciousness. This repeatability among observers can be interpreted as poor to excellent according to the calculated kappa value (Fleiss et al., 2003). Our results show that for most of the indicators, the kappa interpretation strongly varied according to the SH assessed. It happened mainly because kappa values are strongly influenced by the prevalence, and this differed strongly among SHs (when the prevalence is low so is the kappa). The only exceptions to this were in the assessment of spontaneous swallowing and spontaneous blinking where kappa was interpreted as poor agreement among observers in all cases, whereas in vocalisations the kappa was not able to be computed due to lack of outcomes of consciousness. These results suggest that these are cases in which the calculation of kappa does not give much information *per se*. Similarly occurs when paying attention to the PoA found. High PoA may suggest that there is a high agreement among observers. However, it may happen that the agreement is high because the outcome of consciousness of the indicator is very clear to detect for all when present (*e.g.*, presence of head shaking), or because the outcome of consciousness is rarely (*e.g.*, presence of spontaneous blinking and swallowing) or hardly ever observed (*e.g.*, presence of vocalisations). On the other hand, the agreement is lower in the outcomes of consciousness that are more frequently observed (*e.g.*, absence of tonic seizure, presence of breathing).

Inter-observer repeatability of some ABIs for the state of consciousness after WBS in broilers is in general good. The most repeatable indicator before bleeding is vocalisation and spontaneous blinking, followed by tonic seizure and breathing. However, spontaneous blinking and vocalisation was artificially highly repeatable

because hardly ever were observed. When considering these results, we recommend keeping for now tonic seizure and breathing at this stage despite of being less repeatable among observers.

On the other hand, the most repeatable indicators during bleeding are wing flapping, head shaking, and spontaneous swallowing followed by breathing. Nevertheless, spontaneous swallowing and wing flapping were artificially highly repeatable because were observed on few occasions. As the span of observation during bleeding was set from 10 s to 16s distance from neck cutting , sometimes birds start to flap their wings just at the end of this span of time. Thus, it generated doubts at scoring and affected the consensus among observers when wing flapping was present. This reflects the importance of setting the optimal span of observation where greater outcomes of consciousness are observed within a slaughterhouse during bleeding. Additionally, but at lower scale, sometimes there was difficulties at differentiating wing flapping from movements of the wings caused by line shaking. Despite of this, we recommend keeping breathing, wing flapping and head shacking as key ABIs during bleeding despite of being less repeatable.

Repeatability among the three observers could be influenced by impaired visibility towards the animal because of the slaughterhouse design or because when paying attention to a specific ABI, the evaluator is more prone to miss a positive outcome of another ABI. However, it is likely that higher levels of inter-observer reliability could be achieved when standardizing descriptions, training and wider testing at assessing consciousness of broiler at slaughter. Hence, better training looks to be one of the key points to improve animal welfare assessment at slaughterhouse.

4.2. Correlation among animal-based indicators

Based on our observations, pre-stun shocks or runt (small) animals could be responsible of non-stunned birds at the exit of the waterbath due to lack of contact of the head with the electrified water. This explains the presence of broilers that remained conscious and showed combinations of indicators of consciousness before bleeding. On the other hand, some birds did not exhibit tonic seizure and this indicator was not correlated to other outcomes of consciousness before bleeding. It may be possibly caused because the tonic seizure occurred while the bird was submerged in the waterbath as it may happen in long length waterbath or in low line speeds. On the other hand, it is known that when the electrical parameters are set to stun-to-kill the birds, the induction of cardiac arrest leads to reduced or absence of tonic seizure at the exit of the waterbath (EFSA, 2013) and it does not mean that birds are conscious. I this sense, tonic seizure might not be as reliable as the other indicators of consciousness, since it depends on SH configuration and current delivered.

Data on the order of re-appearance of indicators during recovery in poultry is not described in literature. Despite the importance of these indicators, in the context of slaughter, their precise relationships with the brain state or with other indicators of consciousness or unconsciousness are insufficiently known. The study of relationships between different ABIs of state of consciousness may benefit from analyses by correlation (Terlouw *et al.*, 2016b). In the present study, sometimes more than one indicator of consciousness was observed during bleeding. The most observed indicator of consciousness during bleeding was the presence of breathing indicating a return of consciousness. It seems that when a broiler starts breathing is more prone to show movements as head shaking and/or wing flapping so later in the line.

Taking all into consideration, assessing if the broilers are breathing is the most recommended ABI among those assessed in the present study after WBS and before bleeding. However, wing flapping, although not

included, should be also considered before bleeding as some birds suspected to miss the waterbath wing flapped. After bleeding, breathing is the most observed indicator of consciousness and when observed, sometimes was accompanied by wing flapping and/or head shacking.

4.3. Relationship between electrical parameters and stunning efficiency

In the present study, one of the objectives was to compare the effectiveness of stunning between different SH and different key electrical parameters in waterbath. Results showed that some combinations of the electrical parameters were considered effective at inducing and maintaining the state of unconsciousness during bleeding and some did not. It should be highlighted that the wrong choice of electrical parameters or equipment, the use of low voltage/current or/and the application of high frequencies and poor or lack of calibration may represent an animal welfare hazard (Smaldone *et al*., 2021). Thus, precaution when comparing the efficiency of the electrical combinations is needed as the electrical parameters used in waterbath were obtained from the FVO and where not checked *in situ* by researchers in the present study. Additionally, other parameters varied between SHs such as electrode length, distance from head to the electrode, and wetness of shackles that also influence the effectiveness of stunning. Distance from the exit of the waterbath until neck cutting also varied between slaughterhouses so delayed bleeding may increase the prevalence of birds that recover consciousness.

It is known that low frequencies (e.g., 80Hz) are more effective at stunning and require lower current intensities (Girasole et al., 2015). However, different combinations with higher frequencies can also guarantee bird's welfare at slaughtering. In the present study, the prevalence of conscious broilers was calculated and the odds of finding conscious broilers were computed through statistical modelling. Thus, the combination of electrical parameters that resulted in the best effective stunning was found at SH-4 (sine alternative current (AC), 280±57 mA/bird, 358±1 Hz and 64±19 V) were no animal was found with outcomes of consciousness both before and after bleeding despite of delayed bleeding (11 s after stunning). Although conscious broilers at SH-5 (sine AC, 223±1 mA/bird, 352±1 Hz and 205±13 V) were found on 1.6% [0.00-2.00%] birds, the odds of finding birds with outcomes of consciousness differed from that SH-4 where no animals showed outcomes of consciousness. This slight difference may be caused by a higher current applied in SH-4. On the other hand, the odds of finding a conscious bird did not significantly differ between SH-5 (prevalence of 1.6%[0.00-2.00] and applied sine AC, 223±1 mA/bird, 352±1 Hz and 205±13 V) and SH-3 (prevalence of 29.1%[28.4-29.7%] and sine AC 200 mA/bird, 128 to 401 Hz and 160±7 V) and to SH-1 (prevalence of 4.8%[4.9-5.1%] at sine AC, 106±23 mA/bird, 60 Hz, 80V). Although SH-1 applied less current than SH-5, the frequency was much lower and thus, it may lead to similar results.

Electrical combinations that strongly failed at inducing or maintaining unconsciousness were found in most of the batches from SH-2 probably due to the high frequencies (> 600 Hz) applied whatever the current in waterbath (177 to 352 mA/bird). Only one combination gave lower level of conscious birds in this slaughterhouse (batch 5 in SH-2: 5.6%[5.9-6.1%]). Surprisingly, another SH that resulted in high rate of birds recovering consciousness was the SH-6 and no explanation is related to the electrical parameters applied as the low frequency applied (80 Hz) should lead to similar results as SH-1. However, distance from WB to neck cutting in SH-6 was 5 s longer than SH-1 and this may strongly increase the prevalence of birds recovering consciousness. Additionally, it raises the possibility that failure in the maintenance of the equipment or poor or lack of calibration might be also responsible that the electrical parameters recorded differed from what is actually delivered in the WB.

5. Conclusions and recommendations

The most repeatable ABI before bleeding is vocalisation followed by spontaneous blinking. However, both are artificially highly repeatable as hardly ever were observed. On the other hand, absence of tonic seizure is not correlated to other ABIs before bleeding probably because, in certain slaughterhouse, tonic seizure occurred while the bird was submerged. This may happen in long length waterbath or in low line speeds or because the electrical parameters are set to stun-to-kill the birds and the induction of cardiac arrest leads to absence of tonic seizure at the exit of the waterbath and it does not mean that birds are conscious. Therefore, it seems difficult to rely on the absence of tonic seizure to measure consciousness. Finally, we recommend focusing on presence of breathing as indicator of consciousness. However, presence of spontaneous blinking or vocalisation, although hardly ever observed, should not be neglected as indicators of consciousness and ineffective stunning.

During bleeding, the most repeatable ABI was spontaneous swallowing followed by wing flapping, head shacking and breathing. However, spontaneous swallowing is artificially repeatable as was the least observed indicator. Therefore, we recommend focus on presence of breathing, head shaking and wing flapping assessment although less repeatable. Sometimes birds showed simultaneously more than one outcome of consciousness being breathing and head shaking and breathing and wing flapping the most observed combinations.

Repeatability at detecting indicators of consciousness among observers it is likely to be increased by better training and surely is a key point to control animal welfare assessment at slaughterhouse. This work will serve at proposing a refined and validated list of indicators that EURCAW-Poultry-SFA will disseminate for training purposes to Competent Authorities and official inspectors of Member States in the EU. This work will also be available in a factsheet.

The combination of electrical parameters that resulted in a most effective stunning was found when applying 280±57 mA/bird, 358±1 Hz and 64±19 V. Others that kept less than 5% of failure were found when applying less current and similar frequency or both low current and frequency. Combinations that strongly failed at inducing or maintaining the state of consciousness were found when applying high frequencies (> 600 Hz).

6. References

Bates D, Maechler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *J Stat Softw*., 67(1):1-48.

EFSA (European Food Safety Authority). 2013. Scientific Opinion on monitoring procedures at slaughterhouses for poultry. *EFSA Journal,* 11(12):3521.

Gamer M, Lemon J, Singh P. 2019. irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.1. URL: https://CRAN.R-project.org/package=irr

Girasole M, Chirollo C, Ceruso M, Vollano L, Chianese A, Cortesi ML. 2015. Optimization of stunning electrical parameters to improve animal welfare in a poultry slaughterhouse. *Ital J Food Saf*., 4(3):4576.

Landis JR and Koch GG. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33:159–174.

Larsson J. 2020. _eulerr: Area-proportional Euler and Venn diagrams with ellipses_. R package version 6.1.0, URL: https://cran.r-project.org/package=eulerr

Lenth RV. 2021. emmeans: estimated marginal means, aka least-squares means. R package version 1.6.1. URL: https://CRAN.R-project.org/package=emmeans

McHugh ML. 2012. Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3):276-282.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Smaldone G, Capezzuto S, Ambrosio RL, Peruzy MF, Marrone R, Peres G, Anastasio A. 2021. The influence of broilers' body weight on the efficiency of electrical stunning and meat quality under field conditions. *Animals*, 11.

Terlouw C, Bourguet C, Deiss V. 2016a. Consciousness, unconsciousness and death in the context of slaughter. Part I. Neurobiological mechanisms underlying stunning and killing. *Meat Sci.*, 118:133-46.

Terlouw C, Bourguet C, Deiss V. 2016b. Consciousness, unconsciousness and death in the context of slaughter. Part II. Evaluation methods. *Meat Sci*., 118:147-56.