Anim Cogn (2015) 18:361–371
DOI 10.1007/s10071-014-0807-3

ORIGINAL PAPER

Understanding the concept of a reflective surface: Can sheep improve navigational ability through the use of a mirror?

S. D. McBride · N. Perentos · A. J. Morton

Received: 28 May 2014 / Revised: 4 September 2014 / Accepted: 9 September 2014 / Published online: 17 October 2014

Abstract Mirror image-induced stimulation and the ability to use the mirror to improve navigational ability for the purpose of object location are considered measures of animal cognitive ability. The purpose of this study was to assess these cognitive abilities in sheep (Ovis aries) as part of a larger programme profiling the cognitive ability of this animal species. Three separate groups of sheep [($n = 29$); 10 Welsh Mountain, 8 Norfolk Horned and 11 Borderdale] were trained ($\geq$80 % criterion) to locate a salient object (yellow bucket containing cereal-based food) in one of two possible positions, from one of two possible starting points. Each group of sheep was then divided into two sub-groups. One sub-group was exposed to a mirror over a period of 15 days (mirror exposed), whilst the other group remained mirror naïve. All animals were then retested within the choice maze using the mirror, where two out of the possible four bucket positions were now 'apparent' (as reflections in the mirror), in order to assess whether mirror-exposed animals had a more accurate representation of the real bucket position. Sheep exhibited two out of the three archetypal stages of mirror-induced behaviour, namely social/exploratory and contingency behaviour, with differences existing between breeds. Welsh Mountain sheep spent significantly more time fixating on the self-image and touching the self-image with their muzzle than the other two breeds. During the test phase, no overall differences in performance were observed between the mirror-exposed and mirror-naïve groups. However, Welsh Mountain sheep did perform significantly more correct responses overall, compared to the other two breeds. Although the data did not convincingly demonstrate that sheep could use a reflective surface to improve their navigational ability, the observed differences between groups suggests that some breeds of sheep may demonstrate better navigational ability as well as having a greater engagement with the self-image than others.

Keywords Cognition · Spatial representation · Self-awareness

S. D. McBride (✉) · N. Perentos · A. J. Morton
Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Place, Cambridge CB2 3DY, UK
e-mail: sdm47@cam.ac.uk

Introduction

The behavioural reaction of infants and animals to their mirror image has held long-lasting scientific and non-scientific appeal. Pivotal to this attraction is an in-built, although not always clearly defined, concept of 'self' particularly when it is considered that the construct of 'self' does or does not yet exist in the individual (e.g. infants), or may not exist for that particular species. These ideas were formalised by Gallup (1970) and Amsterdam (1972) using primates and infants, respectively. They defined three stages of archetypal mirror-induced stimulation (MIS) behaviour as (1) exploratory and social behaviour as if observing a conspecific, (2) contingency behaviour whereby the individual performs a number of repetitive acts as a way of testing its actions against the visual stimulus and (3) self-directed behaviour whereby the animal investigates a specific part of its body using the mirror, sometimes in the context of a foreign object that is placed upon it (mark test). This response continuum acts as both a marker for human development as well as a demarcation of animal cognitive ability. For example, children under 18 months rarely progress past stage 2 of MIS behaviour

(Bard et al. 2006) and only chimpanzees (Gallup 1970), bottlenose dolphins (Reiss and Marino 2001), the Asian elephant (Plotnik et al. 2010) and Eurasian magpies (Prior et al. 2008) have succeeded in passing the mark test (stage 3).

Mirrors have been used to test goal-orientated behaviour that is not self-directed. For example, a mirror can be used to locate occluded objects if the individual can identify the relationship between the object in the mirror and the object in the real world. Elephants have been observed to locate successfully occluded food using mirrors (Povinelli 1989), as have crows (Medina et al. 2011), monkeys (Anderson 1986; Itakura 1987) and pigs (Broom et al. 2009). It has been argued that objects can be located without the individual having any understanding that either itself or the object viewed in the mirror are representations of the real world (Povinelli 1989; Anderson and Gallup 2011). This argument states that the animal simply uses information contained within the mirror image as a relative spatial cue in order to navigate to the salient object. Nonetheless, these studies still demonstrate the ability of the individual or species to use the generalised concept of reflection in different contexts as a way of improving navigational ability. MIS and mirror use are therefore potentially useful approaches for characterising and comparing the cognitive ability of different animal species. Here, we applied use of mirrors to increase our understanding of the cognitive profile of sheep, using a locomotory-based decision-making approach. This species is currently being developed as a large animal model for human neurological disease (Morton and Howland 2013), and thus, quantifiable measures of cognitive integrity may provide powerful tools to monitor the putative efficacy of therapeutic agents targeted at preventing neurodegeneration. In addition, with the increased use of cognitive markers as inferred measures of emotional state [e.g. Harding et al. (2004)], characterising cognitive attributes of companion and farm animal species may provide a basis for further developing practical measures of animal welfare under different husbandry and management settings.

Methods

Animals

The study used three separate groups of sheep (Welsh Mountain: 10 females (aged 47.8 ± 8.3 months), Norfolk Horned: eight mixed sex (aged 22.3 ± 0.5 months), Borderdale: 11 females (aged 18.4 ± 0.5 months), with each group kept on separate testing locations. Further details of each breed of sheep are presented in Table 1. During the experiment, all animals were kept outdoors in separate groups with free access to water, grazing and a field shelter. Sheep were supplemented with a standard ration of 200-g cereal-based concentrate per day (Dodson and Horrell Ewe and Lamb nuts, Dodson and Horrell, UK). On testing days, the latter was provided as the food reward within the operant task. Six of the Welsh Mountain sheep had been used for a previous cognitive study (Morton and Avanzo 2011). Studies were carried out in accordance with the UK Animals (Scientific Procedures) Act, 1986. No licensed procedures were carried out in the course of these experiments. All animals came from and remained as permanent stock held at the University of Cambridge prior and after the experimental work was carried out.

Training

A locomotory-based decision-making task was considered to be the most ethologically relevant approach for sheep, as a non-dextrous species. Sheep were trained within a simple maze constructed from 1-m high-grated and sheeted metal hurdles (Figs. 1, 2). Animals could not travel through or over hurdles. Animals could see through the grated but not through the sheeted hurdles. During sessions of 10 trials, all sheep were trained to locate a salient object (a yellow bucket containing sheep nuts) in one of two possible positions (1 and 2) from two possible starting points (A and B) within the choice maze design (Fig. 1). Thus, when the animal started from A, the bucket was placed either at position 1 (requiring a left turn out of the gate) or position

Table 1 Description of sheep breeds used in the study

Breed	Description
Welsh Mountain	Prolific upland sheep of small–medium size widely used in the UK as part of the meat production industry. This breed is particularly hardy and can withstand the challenges of being kept on extensive upland areas. This breed is also noted for its ability to survive on poor forage quality. The average mature ewe weight is 35 kg, and there are over 11 million Welsh Mountain sheep currently kept in the UK
Norfolk Horned	Medium-sized lowland but hardy breed currently held in the UK for meat production. The average mature ewe weight is 70 kg. Noted for their ability to survive on poor forage quality. The Norfolk horned sheep are a revived breed and there are currently 2,500 in the UK
Borderdale	Borderdales are a medium to large breed that have been developed in New Zealand using the breeds Blue Faced Leicester and Corriedale. They are predominantly kept on lowlands and used for wool and meat production. The average mature ewe weight is 80–90 kg, and there are currently over half a million Borderdale sheep in New Zealand

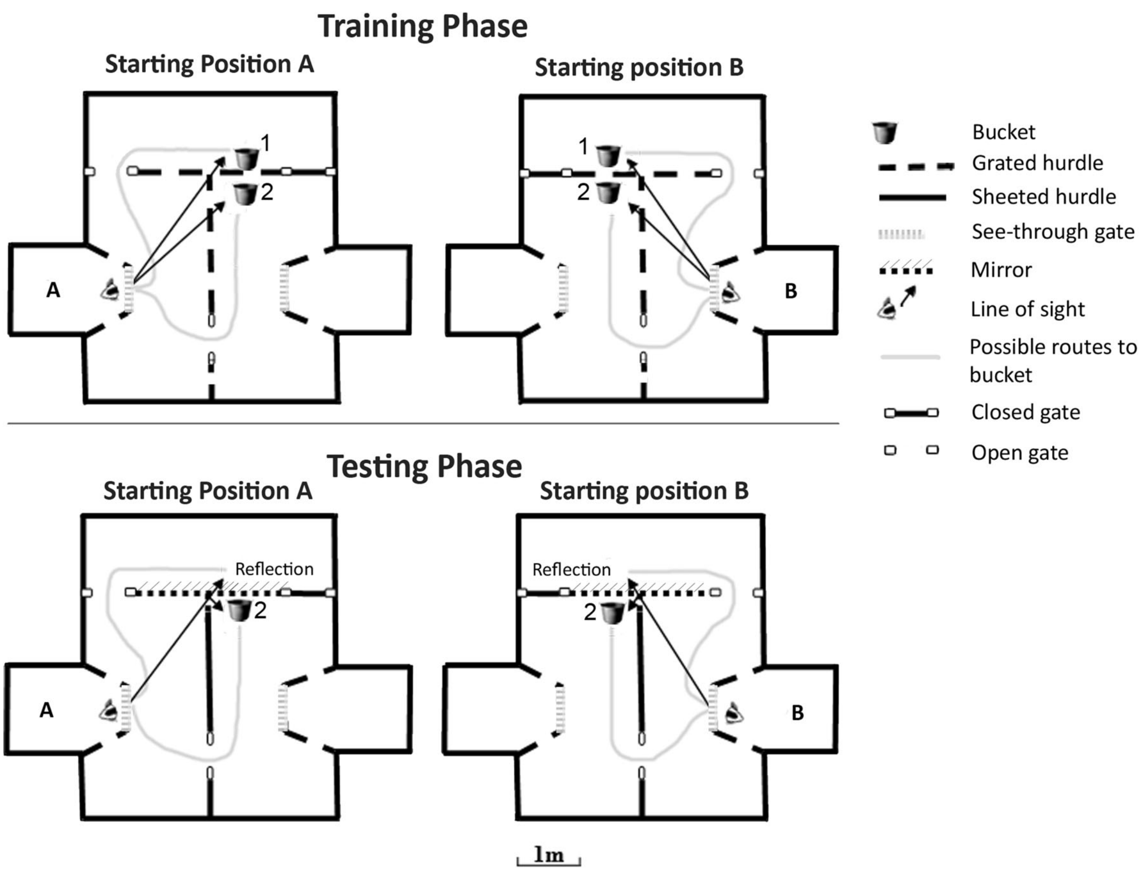
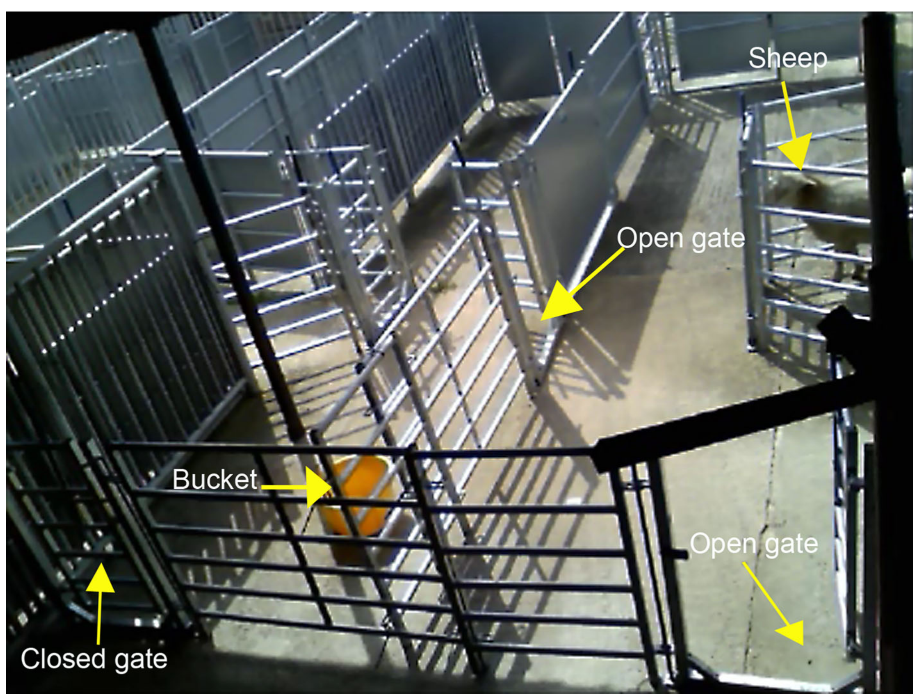


Fig. 1 Layout of apparatus used for the training phase (*upper diagram*) and testing phase (*lower diagram*) during the mirror test. Starting points were *A* and *B* and possible routes to the feed buckets are indicated by the *route lines*. During the testing phase, the grated hurdles were replaced by sheeted hurdles and the mirror

Fig. 2 Photograph illustrating the sheep during the training phase. Starting position is A and bucket position is 2. Note that during the testing phase the grated hurdles are replaced by the mirror and sheet hurdles



2 (requiring a right turn out of the gate). Conversely, when starting from B, the bucket position was placed either at 1 (right turn) or 2 (left turn). The linear distance from the starting point to either bucket was the same. Animals were held at the starting point for 10 s before release whereupon they made the decision to either turn left or right in order to have access to the viewed bucket. The randomised nature of this paradigm was developed such that the decision to go left or right was not side-biased by the outcome of the previous trial, but based solely on the visual information presented at that point in time. This point is critically important in relation to the subsequent testing phase (described in the next section). The learning criterion was set at 80 % (8 out of 10 trials correct) where a correct trial was defined as accessing and eating from the correct bucket. The mean ± SEM number of sessions to reach learning criterion for all animals was 8.3 ± 3.2. Three trainers took part in this experiment; each was responsible for a single cohort of sheep.

Mirror exposure

Each group of sheep was divided into two sub-groups (mirror exposed and mirror naïve), balanced for acquisition rate, previous cognitive test experience (Welsh Mountain) and sex (Norfolk Horned). The mirror-exposed sub-group within each group was then exposed to a perspex mirror (1.2 × 2.4 m, Engineering & Design Plastics Ltd, Cambridge, UK) for 530 min over 15 days in two exposure locations (Figs. 3, 4) following the protocol in Table 2. Animals were initially exposed to the front of the mirror within a confined pen (3 × 4 m) (Area 1). Animals were exposed individually followed by exposure in the presence of a yellow bucket to create a salient point of reference in the reflected image. Animals were then exposed in pairs (with and without the bucket) and then as a group of three. Finally, the whole mirror-exposed group was placed in a larger enclosure (5 × 15 m) (Area 2) where they had access to the front and back of the mirror. Although not the primary objective of the study, animals were also observed during selected sessions (1, 2, 5 and 6) of the mirror exposure to monitor the occurrence of any of the documented stages of behaviour associated with MIS (Table 3). In particular, animals were noted for any contingency behaviour in relation to moving objects (observer and conspecifics) whereby the animal would make visual comparisons through saccade to both the real and reflected object. Fixation time on self-image was measured continuously as a behavioural state (seconds) for the duration of the 5-min sessions. All other behaviours were measured continuously as behavioural counts for the duration of the 5-min sessions. The mirror-naïve sub-groups received a 3-min exposure to the mirror immediately prior to the testing period to reduce novelty effects.

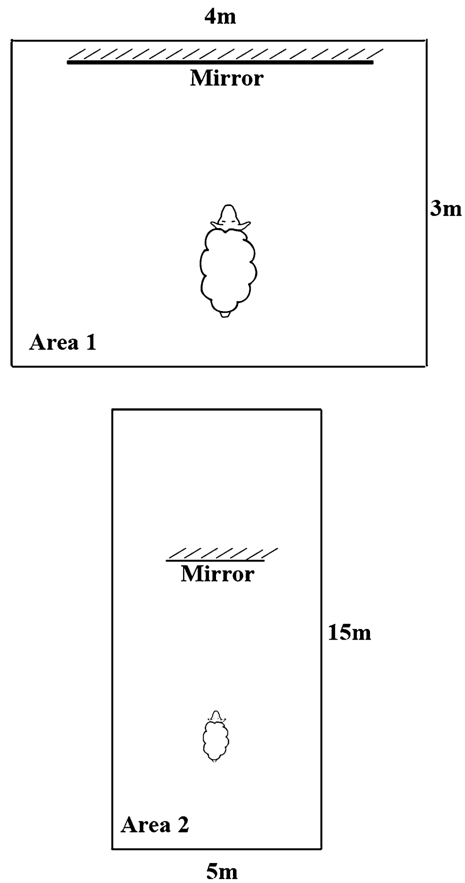


Fig. 3 Layout and arrangement of the two areas of mirror exposure. *Area 1* (3 × 4 m) was for individual and small group mirror exposure, *area 2* (5 × 15 m) was for whole sub-group exposure

Fig. 4 Photograph of sheep observing itself in the mirror in Area 1

Testing

After training and mirror exposure, all animals were retested within a version of the choice maze incorporating

Table 2 Mirror exposure schedule

Day	Type of exposure	Location of exposure	Duration (minutes)
1	Single sheep with observer	3 × 4 m stable; front of mirror only	5
2	Single sheep with observer	3 × 4 m stable; front of mirror only	5
3	Single sheep with yellow bucket and observer	3 × 4 m stable; front of mirror only	5
4	Single sheep with yellow bucket and observer	3 × 4 m stable; front of mirror only	5
5	Two sheep and observer	3 × 4 m stable; front of mirror only	5
6	Two sheep and observer	3 × 4 m stable; front of mirror only	5
7	Two sheep with yellow bucket and observer	3 × 4 m stable; front of mirror only	5
8	Two sheep with yellow bucket and observer	3 × 4 m stable; front of mirror only	5
9	Three sheep and observer	3 × 4 m stable; front of mirror only	5
10	Three sheep and observer	3 × 4 m stable; front of mirror only	5
11	Mirror-exposed sub-group	5 × 15 m area; front and back of mirror	60
12	Mirror-exposed sub-group	5 × 15 m area; front and back of mirror	60
13	Mirror-exposed sub-group	5 × 15 m area; front and back of mirror	120
14	Mirror-exposed sub-group	5 × 15 m area; front and back of mirror	120
15	Mirror-exposed sub-group	5 × 15 m area; front and back of mirror	120

Table 3 Behavioural states and events recorded during mirror sessions 1, 2, 5 and 6, pertaining to the first two documented stages of behavioural response to MIS

Behaviour	Description
MIS Stage 1-Exploratory and social behaviours	
Fixation time on self-image (seconds)	Head directed towards self-image with fixation* on self-image
Touching self-image with nose (counts)	Touching self-image in the mirror with muzzle
Licking (count)	Licking whilst head is directed towards self-image with fixation* on self-image
Attempt to look behind mirror (counts)	Head position and fixation* orientated behind mirror whilst keeping body in front of mirror
Vocalisation (counts)	Vocalisation with head directed towards self-image whilst fixating* on self
MIS Stage 1-Agonistic behaviours	
Stamping (counts)	Stamping with forelimb with head directed towards self-image whilst fixating* on self-image
Ears flattened (counts)	Ears actively flattened in a posterior position whilst fixating* on self-image with head directed towards self-image
Head-butt towards self-image (counts)	Head-butting action towards self-image in the mirror
Barging into mirror (counts)	Barging whole body into the mirror
MIS Stage 2-Contingency behaviour	
Head-movement in relation to image (counts)	Head movements from side to side with head directed towards self-image whilst fixating* on self-image
Fixation on conspecifics (counts)	Saccade to conspecific via the mirror followed immediately by saccade and fixation* on conspecific directly
Fixation on human observer (counts)	Saccade to human observer via the mirror followed immediately by saccade and fixation* on human observer directly

* The lateral placement of the ovine eye confers 310° of monocular and 25°–50° of binocular vision. This extensive peripheral vision is, however, considered to have low acuity and thus sheep will always direct both the head and the eye to bring near and far objects into sharper focus. The term 'Fixation' thus refers to binocular vision with both head and eyes directed the object in question (Piggins and Phillips 1996)

of the mirror (Fig. 1) using a balanced pseudorandom design across the two starting points with approximately half of the treatment group commencing the testing session from either starting point A or starting point B. For both starting points, the bucket was always placed in one position (position 2). Note that with the mirror in place the bucket appeared in position 1. Thus, from both starting points, animals had the same left or right choice as during the training phase, but now during the testing phase, turning left from position A took the animal to an 'apparent' bucket, whereas a right turn took it to the actual bucket. The converse applied when starting from position B. Animals performed 10 trials during the testing phase; first, a probe trial, and then, nine subsequent trials. Since all animals had undergone substantial training during the training phase to locate the bucket from the two positions and (from two random starting locations, the testing phase was developed to be a continuation of this, with the animal making decisions based on the visual information at hand, and thus with minimal bias from the previous trial. It was considered that this approach would allow nine more measurements to be taken after the probe trial with the aim of creating a more complete data set as compared to a single probe trial that is statistically vulnerable to chance [discussed by Gieling et al. (2014)]. In addition, a second yellow bucket containing a food reward was placed behind the mirror to prevent this choice from being extinguished during the course of the 10 consecutive trials. All animals were again held for 10 s for both starting points prior to being released into the maze.

Statistical analysis

Data are mean ± SEM unless otherwise stated.

Inter-rater agreement of coded behavioural data (10 × 5 min sessions) recorded during the mirror exposure was tested statistically using Kendall's coefficient (τ) of concordance.

Behavioural data during the mirror exposure were normalised where necessary to facilitate the use of parametric statistics. Data were then statistically analysed using general linear model to assess the effect of group at each separate time point (Day 1, Day 2, Day 5 and Day 6) and to test for significant changes in behaviour over the measurement period. Tukey's post hoc analysis was used to make individual statistical comparisons between groups and between time-points.

Data from the test session were summarised and analysed in two parts. The first (probe) trial data, recorded as responses to the 'apparent' or 'real' position of the bucket, were analysed using Fisher's exact test to determine whether there was a non-random association between the two categorical variables (correct versus incorrect response) as a result of the treatment effect (mirror exposed versus mirror naïve). Data from the whole session (10 trials) were expressed as the percentage of choices towards the correct ('real') position of the bucket and arcsine transformed for the purpose of data normalisation and facilitating the use of parametric statistical analysis (Zar 1996). The effect of mirror exposure on the whole-session data was tested statistically using a paired T test with group set as the paired blocking factor. The effect of group on the total number of correct responses was also tested statistically using one-way analysis of variance with treatment (mirror exposed, mirror naïve) set as the blocking factor. Post hoc analysis between individual groups was performed using the Tukey's test. In addition, the strength of the primary null hypothesis (H_0: mirror exposure does not affect behaviour within a mirror-based choice test) was tested using Bayesian inference techniques using a bayesian information criterion (BIC) approach as described by Masson (2011). All statistical analysis was carried out using R (version 2.15.2; The R Foundation for Statistical Analysis) in conjunction with Statistica 64 (version 11, 2012; Statsoft Inc.).

Results

Mirror exposure

A significant level of concordance ($T = 0.75$, $P < 0.05$) was recorded between two observers of the coded MIS behavioural data confirming an acceptable level of inter-rater agreement.

During mirror exposure, sheep displayed the first two stages of archetypal MIS responses. A summary of this data is presented in Table 4. One animal (Welsh Mountain) became highly fearful of an uncontrolled noise during testing (building works adjacent to the testing area) during Day 2 and did not complete the test. Behavioural data were not recorded for that animal on that testing day.

Stage 1: Exploratory and social behaviour

All mirror-exposed sheep ($n = 15$) initially spent a considerable amount of time fixating on the self-image (Table 4). For the Welsh Mountain mirror-exposed group, there was also a significant decline of fixation time (minutes) across days (F (3,36) = 4.32, $P = 0.01$; $\eta_p^2 = 0.56$) with Day 1 (191.5 ± 14.4; mean ± SEM) significantly higher than Days 5 (94.3 ± 18.2) and Day 6 (94.3 ± 18.2) (Fig. 5). There was also a significant effect of Group on fixation time (minutes) (F (2,12) = 8.12, $P = 0.006$; $\eta_p^2 = 0.27$) with Welsh Mountain group (191.5 ± 14.4; mean ± SEM) spending significantly more time fixating on the self-image than the Borderdale (36.7 ± 13.3) on Day 1 (Fig. 5). Eight of the mirror-exposed sheep spent time touching the self-image with their nose, and there was a significant effect of time (minutes) (F (3,39) = 3.49, $P = 0.02$; $\eta_p^2 = 0.21$) with Day 1 (12.5 ± 4.8; mean ± -SEM) significantly higher than Day 2 (4.0 ± 2.0), Day 5 (1.3 ± 0.6) and Day 6 (0.7 ± 0.6) for Welsh Mountain sheep (Fig. 6). There was also a significant effect of group on touching behaviour (F (2,13) = 8.03, $P = 0.005$; $\eta_p^2 = 0.55$) with Welsh Mountain (12.5 ± 4.8; mean ± SD) spending significantly more time (minutes)

Table 4 Mean number of times or time spent (minutes) performing mirror-related behaviours during session 1, 2, 5 and 6 of mirror exposure for each group of sheep

Behaviour	Group (number of animals within group performing behaviour)	Day 1 (singles)	Day 2 (singles)	Day 5 (doubles)	Day 6 (doubles)
Stage 1-Exploratory and social behaviour					
Fixation time on self-image (minutes)	Welsh Mountain (5/6)	191.5 ± 14.4	147.0 ± 22.1	94.3 ± 18.2	88.3 ± 16.2
	Norfolk (4/4)	107.8 ± 11.1	53.0 ± 14.8	66.8 ± 23.9	39.3 ± 16.9
	Borderdale (6/6)	36.7 ± 13.3	89.5 ± 20.4	72.2 ± 21.4	59.0 ± 14.1
Vocalisation	Welsh Mountain (3/6)	6.2 ± 2.8	0.8 ± 0.7	0.0	0.0
	Norfolk (1/4)	0.0	0.5 ± 0.4	0.0 ± 0.8	0.0
	Borderdale (1/6)	0.0	0.2 ± 0.2	0.0	0.0
Stamping	Welsh Mountain (1/6)	0.2 ± 0.2	0.0	0.0	0.0
	Norfolk (0/4)	0.0	0.0	0.0	0.0
	Borderdale (0/6)	0.0	0.0	0.0	0.0
Ears flattened (in posterior position)	Welsh Mountain (5/6)	6.0 ± 3.5	1.8 ± 0.9	1.2 ± 0.8	0.3 ± 0.3
	Norfolk (0/4)	0.0	0.0	0.0	0.0
	Borderdale (5/6)	0.2 ± 0.2	1.3 ± 0.7	0.8 ± 0.5	0.3 ± 0.2
Head-butt towards self-image	Welsh Mountain (2/6)	3.2 ± 2.7	0.6 ± 0.5	0.5 ± 0.5	0.8 ± 0.8
	Norfolk (3/4)	0.0	0.0	1.8 ± 0.5	0.0
	Borderdale (1/6)	2.0 ± 1.0	0.0	0.0	1.0 ± 0.9
Touching self-image with nose	Welsh Mountain (5/6)	12.5 ± 4.8	4.0 ± 2.0	1.3 ± 0.6	0.7 ± 0.6
	Norfolk (2/4)	2.0 ± 0.9	0.5 ± 0.5	0.0	0.0
	Borderdale (0/6)	0.0	0.0	0.0	0.0
Vocalisation	Welsh Mountain (3/6)	6.2 ± 2.8	0.8 ± 0.7	0.0	0.0
	Norfolk (1/4)	0.0	0.5 ± 0.4	0.0 ± 0.8	0.0
	Borderdale (1/6)	0.0	0.2 ± 0.2	0.0	0.0
Stamping	Welsh Mountain (1/6)	0.2 ± 0.2	0.0	0.0	0.0
	Norfolk (0/4)	0.0	0.0	0.0	0.0
	Borderdale (0/6)	0.0	0.0	0.0	0.0
Barging into mirror	Welsh Mountain (2/6)	1.0 ± 0.6	0.0	0.0	0.0
	Norfolk (0/4)	0.0	0.0	0.0	0.0
	Borderdale (0/6)	0.0	0.0	0.0	0.0
Licking	Welsh Mountain (4/6)	0.3 ± 0.2	5.0 ± 1.8	0.0	0.0
	Norfolk (0/4)	0.0	0.0	0.0	0.0
	Borderdale (0/6)	0.0	0.0	0.0	0.0
Stage 2-Contingency behaviour					
Head-movement in relation to image	Welsh Mountain (3/6)	1.8 ± 0.7	0.4 ± 0.3	1.7 ± 1.1	0.5 ± 0.3
	Norfolk (2/4)	1.3 ± 0.5	0.0	0.0	0.0
	Borderdale (2/6)	0.8 ± 0.8	1.0 ± 0.7	0.5 ± 0.5	0.0
Fixation on conspecifics	Welsh Mountain (5/6)	–	–	2.3 ± 0.7	2.2 ± 0.4
	Norfolk (3/4)	–	–	1.3 ± 0.5	1.3 ± 0.9
	Borderdale (5/6)	–	–	1.7	0.7 ± 0.4
Fixation on human observer	Welsh Mountain (2/6)	0.0	0.2 ± 0.2	0.2 ± 0.2	0.0
	Norfolk (0/4)	0.0	0.0	0.0	0.0
	Borderdale (0/6)	0.0	0.0	0.0	0.0
Attempt to look behind mirror	Welsh Mountain (2/6)	0.2 ± 0.2	0.2 ± 0.2	0.3 ± 0.3	0.0
	Norfolk (1/4)	0.0	0.3 ± 0.2	0.0	0.0
	Borderdale (0/6)	0.0	0.0	0.0	0.0

touching the self-image than the Norfolk Horned (2.0 ± 0.9) and the Borderdale (0.0 ± 0.0) on Day 1 (Fig. 6).

Welsh Mountain and one Norfolk Horned sheep attempted to look behind the mirror. An aggressive/fearful posture of flattened ears was performed by all of the Welsh

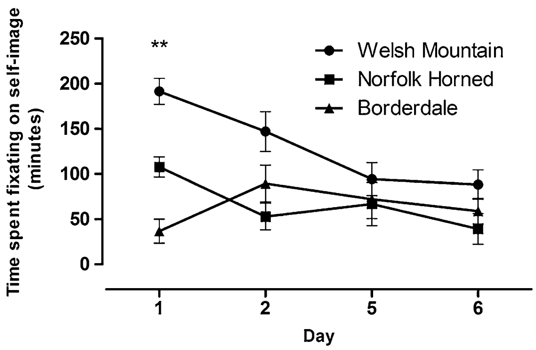


Fig. 5 The time spent fixating on self-image for three breeds of sheep. Data were collected during four sessions (5 min) of mirror exposure (Day 1, Day 2, Day 5 and Day 6) for three groups of sheep. $^{**}P < 0.01$

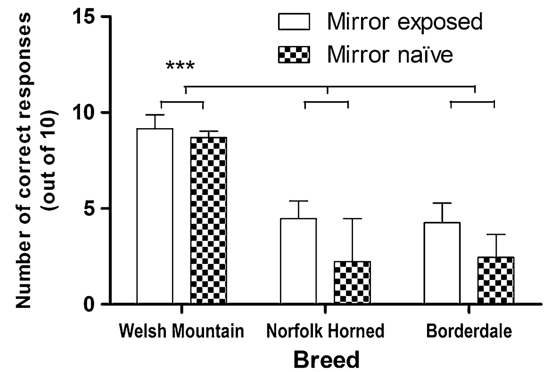


Fig. 7 The number of correct responses for three breeds of sheep (Welsh Mountain, Norfolk Horned, Borderdale) during the choice maze task.$^{***}P < 0.001$

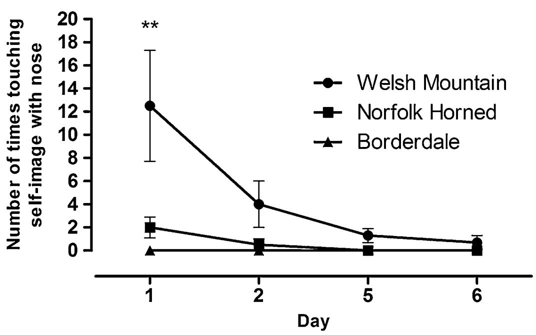


Fig. 6 The number of times touching self-image with nose for three breeds of sheep. Data were collected during four sessions (5 min) of mirror exposure (Day 1, Day 2, Day 5 and Day 6). $^{**}P < 0.01$

Mountain and Norfolk Horned sheep towards the mirror image but by none of Borderdale (Table 4). Six individuals (2 × Welsh Mountain, 3 × Norfolk Horned, 1 × Borderdale) also physically assaulted the mirror image by head-butting, and two of the Welsh Mountain sheep also barged their bodies sideways into the mirror (Table 4). Five animals vocalised during testing (3 × Welsh Mountain, 1 × Norfolk Horned, 1 × Borderdale), one animal stamped (Welsh Mountain) and four animals licked (Welsh Mountain) during the observed sessions (Table 4).

Stage 2: Contingency behaviour

During pair exposure, thirteen individuals (5 × Welsh Mountain, 3 × Norfolk Horned, 5 × Borderdale) fixated on conspecifics via the mirror and then saccaded immediately to the actual animal (Table 4). Two of the Welsh Mountain sheep fixated on the observer via the mirror. Seven individuals (3 × Welsh Mountain, 2 × Norfolk Horned, 2 × Borderdale) were observed to perform contingency-type behaviour, observing and testing their own movement in the mirror.

Choice maze task

During the probe trial (Trial 1), there was no significant difference (Fisher exact test; $P = 0.64$) in the number of correct responses (to the real bucket position) between mirror-exposed to mirror-naïve animals (six versus three animals responding correctly, respectively). Statistical analysis of the correct number of responses for the whole of the test session also demonstrated no significant difference ($t\,(13) = -0.82$, $P = 0.42$) between the mirror-exposed and the mirror-naïve groups (4.7 ± 1.0 versus 3.9 ± 1.0). The probability of the null hypothesis being true (H_0: mirror exposure does not affect behaviour within a mirror-based choice test), given the observed data set, was strong (as defined by Raftery 1995) at $P = 0.84$. A significant difference was observed, however, between groups (mirror-exposed and mirror-naïve animals combined) with the Welsh Mountain sheep performing significantly more correct responses compared to the Norfolk Horned and Borderdale groups ($F(2,24) = 22.2$, $P < 0.001$; $\eta_p^2 = 0.68$; 95 %) (Fig. 7). The 95 % confidence intervals for this test were as follows: Welsh Mountain [8.2, 9.7], Norfolk Horned [1.0, 5.7] and Borderdale [1.8, 4.8].

Discussion

The behavioural measures recorded during the exposure to the mirror closely resemble those that have been observed in other species when exposed to reflective surfaces, referred to as MIS. As previously stated, this was first reported by Gallup (1970) and Amsterdam (1972) and typically follows three sequential stages of behaviour (1) exploratory and social behaviour, (2) contingency behaviour and (3) self-directed behaviour. All sheep in this study demonstrated the first of these behavioural stages with eight animals touching the mirror image with their nose and all sheep fixating on their image for prolonged periods of time during the mirror exposure. Several of the sheep also flattened their ears in the posterior position in a manner commonly exhibited in this species as an aggressive posturing (Lynch et al. 1992) or associated with unfamiliar and unpleasant uncontrollable situations (Boissy et al. 2011). Six animals also physically assaulted the mirror through the standard ovine agonistic response of head-butting and barging, with one animal performed stamping behaviour (Lynch et al.1992). Submissive posturing was also exhibited, with four animals performing licking behaviour. Although flocking behaviour in sheep has been extensively researched (Lawrence and Woodgush 1988; Michelena et al. 2008, 2009; King et al. 2012), the number of studies that have quantified individual ovine social behaviour has been more limited. One study, Pokorna et al. (2013), whilst comparing time-budgets between sheep ($n = 60$) and goats ($n = 20$) on grassland, presented a mean value of 1.6 % of time spent on social interactions. This value seems low for a social species but must be taken in the context of other predominating feeding behaviours such as grazing and ruminating, reported within the same study at 71 %. Social responses reported within the present study during mirror exposure support the notion that sheep are a highly gregarious species. More importantly, they strongly suggest that even a mirror image will elicit social behaviours where the animal is perceiving the mirror image as an unknown conspecific.

After the second session of mirror exposure, contingency-type behaviours started to become apparent, with some animals fixating on the conspecific mirror image and then saccading to the actual flock member. Two of the Welsh Mountain sheep also did this with the familiar human observer. Several of the animals for all three groups also moved their head back and forward in a repetitive manner which is interpreted as archetypal contingency behaviour whereby the animal is correlating its movements with changes in the visual image (Plotnik et al. 2010). Three animals also appeared to look behind the mirror although it was difficult to dissociate this from general exploratory actions. Overall, whilst all animals performed the first stage of MIS behaviour (exploration), only some (7 out of 29) appeared to proceed to certain types of contingency behaviours. This type of individual variation in responses to MIS has also been observed in other species (Prior et al. 2008; Plotnik et al. 2010).

Interestingly, Welsh Mountain sheep were significantly different in some of their responses to the mirror from sheep of the other two breeds, with the most notable being the significantly higher proportion of time spent fixating (Fig. 5) and touching (Fig. 6) the self-image. This could be inferred as differences in exploratory or social attributes of this particular group.

During the choice maze task, no significant difference was observed between breeds in their ability to locate the 'real' position of the bucket either during the first probe trial or over the course of the 10 trails. This suggested that sheep overall were not applying the concept of a reflective surface in order to improve navigational ability. A major caveat to this conclusion, however, is that both sub-groups of the Welsh Mountain sheep (mirror exposed and mirror naïve) predominantly chose the correct (actual) bucket position compared to the other two groups (8.9 ± 0.4 versus 3.4 ± 1.1 and 3.6 ± 0.9) over the course of the 10 trials in the testing session (Fig. 7). Moreover, given the confidence intervals for each group, the Welsh Mountain group's correct responses [8.2, 9.7] were significantly above chance (50 %), whereas the Norfolk Horned [1.0, 5.7] and Borderdale [1.8, 4.8] sheep were either near significance or significantly lower than chance, suggesting that the latter two groups were actively choosing the apparent, as opposed to the real bucket position.

There are potentially a number of different reasons for this difference in results. For example, although every possible step was taken to replicate exactly the design of the experiment between the two testing sites, through detailed measurement of distances and angles between primary structures within the testing area, it is possible that slight differences may have existed. This could have affected, for example, the perception of distance between the starting point and the visual location of the buckets which, in turn, may have biased the decision-making process. The second explanation may relate to a trainer effect, that in some way, trainer 1, who was responsible for the Welsh Mountain sheep, influenced the choice at the point of testing such that the actual bucket position was chosen predominantly, although all trainers worked from the same standardised and detailed training protocol. A recent study attempting to replicate the operant use of mirrors in pigs by Broom et al. (2009) using exactly the same experimental design and protocol was unable to do so (Gieling et al. 2014). Although this may simply reflect a Type I statistical error, it may also potentially demonstrate the subtle but important differences that exist when testing different cohorts of animals. The latter study also identified the degree of engagement with the mirror as a determining factor of operant mirror use, and indeed, it was the Welsh Mountain sheep in this study that demonstrated significantly more time performing mirror-orientated behaviour during mirror exposure. A third explanation for the difference between groups may pertain to phenotypic differences, in that either genotype, life experience or the additive effect of both, allowed both the mirror-exposed and the mirror-naïve animals within the Welsh Mountain group a more accurate interpretation of the reflective surface. More specifically, these animals were able to inhibit the prepotent visual stimuli of the apparent bucket location and use knowledge about the mirror (either spatially or as a navigational cue) to locate the real position of the bucket. The Welsh Mountain sheep were on average 20.4 months older than animals within the other two groups although all animals were developmentally mature (over 18 months) and none of the females had reproduced. Thus, it is unlikely that developmental milestones were a contributing factor in explaining group difference. In contrast, these older animals may have had additional prior exposure to reflective surfaces, either out in field conditions (e.g. pooled water) or as a result of previous operant testing. As previously stated, the majority of this group of animals had already undergone some operant testing and indeed much of this involved the use of sheeted hurdles which can, under certain conditions, have reflective characteristics. The differences in results may also say something about genetic differences in spatial navigation. For example, even though some animals have the ability to process correctly the additional spatial information that a reflective surface confers, they may still have difficulty in utilising this information for navigational purposes. Route determination may be more established in the Welsh Mountain sheep compared to the other two breeds due to the more extensive farming methods under which these animals are generally managed. Welsh Mountain sheep are kept on expansive and unfenced upland areas, whereas the Norfolk Horned and Borderdale, as lowland breeds, tend to be kept in smaller fenced fields. Thus, the former are renowned for self-determined movement over much larger areas to access grazing, water and shelter, referred to as hefting (Morton and Avanzo 2011). In this sense, enhanced spatial navigation is a critical feature for their survival. Thus, even though the three breeds in this experiment were kept under general lowland husbandry management, the different selection pressures placed upon Welsh Mountain sheep as a breed may mean that they have enhanced navigational ability. Whether this reflects enhanced spatial mapping attributes or a greater ability to use proximal cues during route navigation is something to be established in future research studies, as is the more general effect of domestic selection and environment on animal cognition.

Conclusion

Overall, our data did not convincingly demonstrate that sheep could use a reflective surface to produce more informed decisions about route navigation within an operant context. The behaviour expressed during mirror exposure suggested that some sheep may have the ability to form a concept of self but we were not able to demonstrate this conclusively. Using operant paradigms to ask questions about conscious appreciation of self is intrinsically difficult due to the various interpretations of results (Medina et al. 2011). Nevertheless, the potentially interesting behavioural results observed during the mirror exposure warrants further investigation into this area, given that no previous research exists for this species.

Finally, interesting differences were observed between groups, with Welsh Mountain sheep potentially showing higher functioning than the other breeds with regard to route navigation as well as a greater engagement with the self-image. This could be interpreted as a greater understanding of 'self', breed differences in spatial navigation ability or the effect of operant task experience. This study provides a relevant contribution to the knowledge base on the cognitive ability and limitations of an ungulate species that is becoming increasingly used as a model for cognitive function.

Acknowledgments We thank Calum Frost for his invaluable work training and testing the sheep. This work was funded by CHDI Inc.

References

Amsterdam B (1972) Mirror self-image reactions before the age 2. Dev Psychobiol 5:297–305

Anderson JR (1986) Mirror-mediated finding of hidden food by monkeys (*Macaca tonkeana* and *Macaca fascicularis*). J Comp Psychol 100:237–242

Anderson JR, Gallup GG (2011) Do rhesus monkeys recognize themselves in mirrors? Am J Primatol 73:603–606

Bard KA, Todd BK, Bernier C, Love J, Leavens DA (2006) Self-awareness in human and chimpanzee infants: what is measured and what is meant by the mark and mirror test? Infancy 9:191–219

Boissy A, Aubert A, Desire L, Greiveldinger L, Delval E, Veissier I (2011) Cognitive sciences to relate ear postures to emotions in sheep. Anim Welfare 20(1):47–56

Broom DM, Sena H, Moynihan KL (2009) Pigs learn what a mirror image represents and use it to obtain information. Anim Behav 78:1037–1041

Gallup GG (1970) Chimpanzees: self-recognition. Science 167:86–87

Gieling ET, Mijdam E, van der Staay FJ, Bordquist RE (2014) Lack of mirror use by pigs to locate food. Appl Anim Behav Sci 154:22–30

Harding EJ, Paul ES, Mendl M (2004) Animal behavior: cognitive bias and affective state. Nature 427(6972):312

Itakura S (1987) Mirror guided behavior in Japanese monkeys (Muscata-Fuscata-fuscata). Primates 28:149–161

King AJ, Wilson AM, Wilshin SD, Lowe J, Haddadi H, Hailes S, Morton AJ (2012) Selfish-herd behaviour of sheep under threat. Curr Biol 22:R561–R562

Lawrence AB, Woodgush DGM (1988) Home-range behavior and social-organization of Scottish Blackface sheep. J Appl Ecol 25:25–40

Lynch JJ, Hinch GN, Adams DB (1992) The behaviour of sheep: biological principles and implications for production. C.A.B. International, Wallingford

Masson MEJ (2011) A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. Behav Res Methods 43(3):679–690

Medina FS, Taylor AH, Hunt GR, Gray RD (2011) New Caledonian crows' responses to mirrors. Anim Behav 82:981–993

Michelena P, Gautrais J, Gerard JF, Bon R, Deneubourg JL (2008) Social cohesion in groups of sheep: effect of activity level, sex composition and group size. Appl Anim Behav Sci 112:81–93

Michelena P, Sibbald AM, Erhard HW, McLeod JE (2009) Effects of group size and personality on social foraging: the distribution of sheep across patches. Behav Ecol 20:145–152

Morton AJ, Avanzo L (2011) Executive decision-making in the domestic sheep. PLoS ONE 6:e15752

Morton AJ, Howland DS (2013) Large animal models of Huntington's disease. J Hunt Dis 2:3–12

Piggins D, Phillips CJC (1996) The eye of the domesticated sheep with implications for vision. Anim Sci 62:301–308

Plotnik JM, de Waal FBM, Moore D, Reiss D (2010) Self-recognition in the Asian elephant and future directions for cognitive research with elephants in zoological settings. Zoo Biol 29:179–191

Pokorna P, Hejcmanova P, Hejcman M, Pavlu V (2013) Activity time budget patterns of sheep and goats co-grazing on semi-natural species-rich dry grassland. Czech J Anim Sci 58:208–216

Povinelli DJ (1989) Failure to find self-recognition in Asian elephants (elephas-maximus) in contrast to their use of mirror cues to discover hidden food. J Comp Psychol 103:122–131

Prior H, Schwarz A, Guentuerkuen O (2008) Mirror-induced behavior in the magpie (Pica pica): evidence of self-recognition. PLoS Biol 6:1642–1650

Raftery AE (1995) Bayesian model selection in social research. Sociol Method 25:111–163

Reiss D, Marino L (2001) Mirror self-recognition in the bottlenose dolphin: a case of cognitive convergence. Proc Nat Acad Sci USA Stat Am 98:5937–5942

Zar JH (1996) Biostatistical analysis, 3rd edn. Prentice Hall, Upper Saddle River