Towards testing of identity concepts in horses: conditional discrimination learning*

Björn Samuelsson
bjorn.samuelsson@gmail.com

June 5, 2008

Experiments have shown that horses are capable of some types of discrimination learning and that they can learn the concept of relative size, but it is unknown whether horses can select a test stimulus by similarity to a sample stimulus. In this project, four horses failed to learn a two choice matching-to-sample task without a mechanism that ensured that they paid attention to the sample stimulus. Learning performances for conditional discrimination tasks, with similarities to the original matching-to-sample task, were also investigated and one horse was successful in learning such a task. These results point out interesting aspects of the horses' learning behaviors, but more research is needed to investigate the validity of the observed trends.

1 INTRODUCTION

Despite a vast amount of literature about horse training, there is little scientifically verified knowledge about learning behavior in horses. See Murphy and Arkins (2007) for a review of research about learning behavior in horses. In particular, there is little research about horses' abilities of concept formation and relational learning. However, some researchers have found evidence for learning processes that can be seen as indicators of relatively advanced inference abilities in horses. Hanggi (1999) demonstrated that horses can learn to select symbols with an open center in two-choice discrimination tasks where one stimulus had an open center and the other had a filled center. After the horses had learned this task for a small set of stimuli, they carried out the same task for novel stimuli with a performance that was significantly above chance. Hanggi (2003) investigated discrimination based on relative size and showed that this ability was transferred to novel symbols and from symbols to solid objects. Furthermore, the investigations by Roth *et al.* (2007) show that horses can select colored stimuli based on relative hue. Flannery (1997) addressed whether horses can use a concept of visual identity. However, the task that the horses learned can be carried out with a strategy based on direct associations between symbols and rewards. Thus, the results presented by Flannery (1997) can be explained without the usage of a general concept of visual similarity.

In this thesis, I tested to what extent horses learned a matching-to-sample task, and I also tested the learning performance of the horses in two other tasks that are related to the first one. All tasks involved conditional discrimination and they cannot be solved without the information from a sample stimulus that determined which one of two test stimuli (or locations) that was rewarded. The time frame for my experiments turned out to be too narrow to test identity-related concept formation using my experimental approach. However, the results can be useful for future research on such concept formation in horses.

For other species, identity-related concept formation has been investigated by learning and testing of *identity matching-to-sample* tasks. In a matching-to-sample (MTS) task, a sample stimulus is shown to the subject that then selects one of two or more test stimuli. Only one of the test stimuli, the so called target stimulus, is rewarded. A basic feature of matching-to-sample tasks is that the sample determines which test stimulus is rewarded. This means that the subject cannot simply associate a set of test stimuli with a reward if the experiment is carried out properly — each test stimulus is equally likely to be rewarding if the information from the sample is disregarded. The term "identity matching-to-sample" (IMTS) refers to tasks where the target stimuli is identical to the sample stimulus in some chosen aspect. In many IMTS experiments, the aspect to be matched is overall visual appearance but other aspects (e.g., scent) can also be used. (Typical MTS tasks are based on identity, but other forms of MTS include matching of abstract properties and learning of arbitrary matching relations.)

Learning of an IMTS task with a limited set of available stimuli is not sufficient to demonstrate formation of a similarity concept — introduction of novel stimuli is needed to probe to what extent the learned behavior relies on a general concept. The ability to match previously unseen stimuli has been demonstrated for apes (see, e.g., Vonk, 2003; Oden *et al.*, 1988), monkeys (see, e.g., D'Amato *et al.*, 1985) and pigeons (see, e.g., Lombardi, 2008; Wright, 2001). Similar findings have also been presented for marine mammals like seals (Mauck and Dehnhardt, 2005). Rats have also been tested in matching-to-sample experiments, and in an experiment using scent stimuli instead of visual stimuli Peña *et al.* (2006) obtained strong evidence for a concept of equal scents.

Given the variety of species that are capable of learning IMTS and to generalize the matching rule to new stimuli, it seems likely that horses have similar capabilities. Because horses have been able to learn categorization and size comparison of symbols, visual IMTS might be suitable for testing similarity-based concept

*Master thesis in cognitive science supervised by Christian Balkenius formation. However, it is not straightforward to ensure that the horse pays attention to the sample stimulus. In a typical MTS setting, only the sample stimulus is shown initially and the subject has to perform some action (like pecking on or touching the sample stimulus) for the test stimuli to appear. To require the horse to touch the sample stimulus with its muzzle might have drawbacks because the visual perception of the stimulus is likely to be compromised by its proximity when the horse is touching it.

Wright (2001) investigated the role of pecking on the sample stimulus for pigeons learning a matching-to-sample task. Pigeons that did not need to peck on the sample or to peck on it only once learned more slowly than pigeons that were required to peck 10 or 20 times before the test stimuli were displayed. Furthermore, the generalization performance depended strongly on the number of pecks. The group that pecked 20 times displayed full generalization to new stimulus configurations and to new stimuli, whereas the groups that pecked 0 or 1 times apparently learned the individual stimulus configurations rather than the matching rule. The performance of the pigeons that pecked 10 times indicated partial concept learning.

Similar effects of responding to the sample stimulus have also been demonstrated for monkeys. Katz *et al.* (2002) used a same/different learning task where one sample picture and one test picture were displayed along with a white rectangle. If the sample and the test pictures were identical, touching the test picture was rewarded. Otherwise, touching the white rectangle was rewarded. Monkeys that did not need to activate each trial by touching the sample stimulus typically failed to learn the task, whereas all monkeys that touched the sample stimulus 10 times before the test stimulus appeared learned the task. The latter group also displayed generalization to new stimuli. The generalization performance in a same/different task is also strongly dependent on the size of the training set both for monkeys and for pigeons, as demonstrated by Wright and Katz (2006). The performance with new stimuli were as good as the performance with known stimuli for sufficiently large training sets. The required set size varied from 64 to 256 items between the animals with the monkeys learning somewhat faster than the pigeons.

With a suitable experimental setup, the learning rate can be boosted dramatically. Wright and Delius (1994) trained pigeons to dig for seeds buried in colored gravel and tested performance in matching-to-sample tasks and oddity-from-sample tasks. Oddity-from-sample (or nonmatching-to-sample) works along the same lines as a matching-to-sample but the nonmatching test stimulus is rewarded rather than the matching one. In the experiment by Wright and Delius (1994), the pigeons learned the task for black and white gravel after approximately 30 trials. In experiments where the pigeons only peck on the sample stimulus, hundreds or thousands of trials are usually required for learning to occur. Seemingly small differences in the experimental setup can also be important for the outcome. Lombardi (2008) investigated generalization of matching-to-sample and oddity-from-sample for pigeons in a situation where the stimuli were presented on lids on either side of the sample stimulus. Generalization did occur when the lids were separated by 1 cm from the sample but not when there were no such separation. Furthermore, Lombardi (2008) demonstrated that the pigeons were capable of generalizing the matching rule from matching of different colors to matching different of black shapes presented on a white background.

One way to test for a concept of similarity without the problems associated with attention to a sample stimulus is to present several test stimuli where one stimulus is different from the other stimuli. The odd stimulus is rewarding and such an experiment is referred to as an “oddity task”. In contrast to oddity-from-sample tasks, there is no sample stimulus in oddity tasks. In an experiment by Roitberg and Franz (2004), 17 goats were trained in an oddity task. Only one goat learned the task, but the failure of the other goats to learn the task might not be caused by inability to learn an oddity task. As pointed out by the authors, the fact that the goats could make an unlimited number of attempts immediately after a failed trial might favor strategies that are based on persistent responding until the goat randomly chooses the rewarding alternative.

The stimuli setup in this project is an extension of the experimental setups used by Macuda and Timney (1999), Geisbauer *et al.* (2004), and Roth *et al.* (2007) with test stimuli displayed on small doors with rewards placed behind them. In this project, a sample stimulus was added and it was displayed at the center of the equipment. Configurations with a centrally displayed sample stimulus are typical for MTS-experiments but apparently untested for horses in prior research. I tested such a configuration in the context of IMTS and conditional discrimination.

2 MATERIALS AND METHODS

2.1 Experimental design

The trials were two-choice tasks where one choice was rewarded depending on a sample stimulus. Carrot pieces were placed on a shelf behind two 30 cm wide square doors in a plywood wall. The doors could be locked and their hinges were located at the top of the doors so each door was closed by its own weight. Wooden pieces running along the rear sides of the doors prevented the doors from swinging forward and ensured that the locking mechanism was not visible from the front side of the wall.

The plywood wall (see figure 1) is 2 m wide, 1 m tall, and painted light gray. The distance from the ground to the bottom of the wall can be set to 30, 40, 50, or 60 centimeters. At the middle of the wall, there is a protruding holder for a sample stimulus. It extends 30 cm from the wall, is 40 cm wide, and stretches from the bottom of the wall to its top. In the initial experimental setting, symbolic stimuli were used. For such trials,

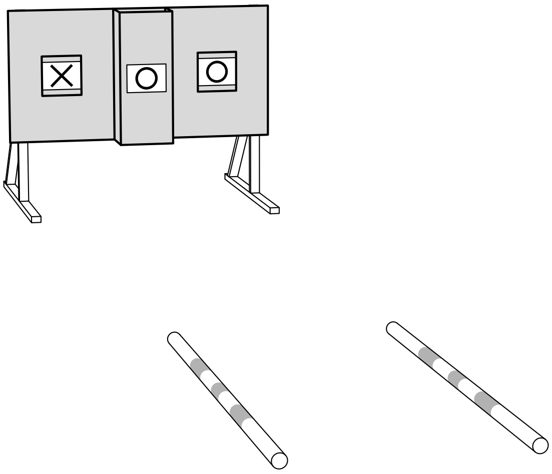

Figure 1. Experimental apparatus for matching-to-sample trials. The sample stimulus is placed in a holder at the middle of the apparatus and the test stimuli are placed on doors on both sides. There is a piece of carrot behind each door, and the door with the matching stimulus is unlocked whereas the other door is locked. At the start of each trial, the horse is positioned between parallel poles placed on the ground. The distance from the poles to the plywood wall of the apparatus is 5 m.

the stimuli were printed on A4-papers and mounted on the central holder and on each door.

In a matching-to-sample trial, one piece of carrot was placed behind each door. One door displayed a symbol that was identical to the sample symbol, whereas the other door displayed a different symbol. The door with the matching symbol was unlocked and the other door was locked. Before each trial, the horse was positioned 5 m in front of the wall. This distance was marked by the ends of two poles that were placed on the ground (see figure 1). After the horse was allowed to approach the wall, the first attempt to push a door open was registered as a correct or incorrect response depending on the selected door. In addition to the piece of carrot, a correct response was rewarded with praise and patting. The sequences of stimuli were generated independently and each sequence was only used once. This procedure enabled blinded trials where the horse handler did not know the selected sequence. The handler also avoided to look at the stimuli before the horse had reached the stimuli setup.

2.2 Subjects

The subjects in this thesis are four experimentally naive American Quarter Horses: Elna (8 year mare), Diamond (22 year mare), and Monica (4 year mare), and Bruno (7 year gelding).

2.3 Shaping

In all sessions, the plywood wall was mounted with its lower edge 50 cm above the ground, placing the centers of the doors at a level of 100 cm. The first four days, with one training session a day for each horse, were devoted to habituation to the equipment and teaching the horses to approach the wall and open a door to retrieve a reward. The horses were initially led to the wall with halter and lead rope and the doors were first opened manually. Directly after the horses had learned to open a partially open door, they did not locate a closed door until it was nudged somewhat. On the fourth day, all horses succeeded in opening closed doors and this day was also the first day when the horses were turned loose in front of the wall. The halter was removed at the first loose trial and a rope resting on the neck of the horse was used to guide the way back after each trial into position for the next trial.

2.4 Initial matching-to-sample training

Three different symbols (O, ×, and ⊥) were used as stimuli in the initial training. The initial training is divided into phases where the number of trials per session were 6, 10, and 20 for the first, second, and third phase, respectively. In the first phase, with 6 trials per session, a trial was reiterated after each incorrect response, meaning that the horse was led back to the starting position and the trial was repeated until the door with the matching stimulus was selected. In the second phase, the horse was allowed to make additional attempts only for the last 2 trials and only if no more than two responses in the first 8 trials were correct. In the third phase, additional attempts were limited to the last stimulus configuration in a session.

There are 6 different ways to select symbols for the sample stimulus and the distractor (the incorrect test stimulus). In the first phase, the sequence of symbol selections was set to a random permutation of the 6 possible combinations. To fully determine a stimulus configuration, a position (left (L) or right (R)) for the matching test stimulus has to be chosen. The sequence of positions for each session was chosen in accordance with a string randomly selected from the set {LLRLRR, LLRRLR, LRLLRR, LRLRRL, LRRLLR, LRRLRL, RLLRLR, RLLRRL, RLRLLR, RLRRLL, RRLLRL, RRLRLL}. This set of strings balances correct responses between left and right, ensures that the same side is not correct for more than two consecutive trials, and it excludes the purely alternating sequences LRLRLR and RLRLRL.

In the second and third phases, sequences of stimulus configurations were created with each configuration randomized uniformly among all 12 possible configurations. Sequences were checked with respect to reasonable balance in the positions of the matching test stimuli for each session. If the number of occurrences of matching test stimuli appearing at a given side was less than a cutoff c, the sequence was discarded. I set $c = 3$ for the 10-trial sessions and $c = 7$ for the 20-trial sessions. Throughout the experiments, there were typically 3–5 experimental sessions per week, but some weeks contained 0–2 sessions.

In Monica's ninth session in the first phase, the handler started to follow her along towards the stimuli without releasing the rope. This change to the procedure was introduced due to an apparent lack of focus on the task. As the second phase started, the approach to the wall was altered in the same way for the other horses, because they tended to lower their heads and did not pay attention to the sample stimulus. The handler avoided giving inadvertent cues about the correct choice by looking away from the stimuli while walking towards the wall. (As pointed out above, the procedure is setup in such a way that the handler did not know the stimulus configuration beforehand.)

2.5 Object-to-position matching

The matching-to-sample training was interrupted for Elna and Bruno and replaced by an apparently simpler task, because these horses appeared to pay too little attention to the sample stimulus in the original setting. This change was carried out 50 days after the first session of habituation/shaping, giving a total number of 220 and 210 MTS-trials for Elna and Bruno, respectively. In the new task, the symbolic stimuli were removed and for each trial an object was placed on the ground approximately half a meter in front of the sample stimulus holder. The object was either a small cone or a foam block for cavaletti obstacles. (A cavaletti obstacle is like a small jumping obstacle with poles on the ground or a few decimeters above the ground.) The height of the cone was 30 cm and the size of the foam block was 40 cm × 30 cm × 9 cm. For Bruno, the correct response was pushing the door on the left side if the sample was the cone and the door on the right side if the sample was the block. For Elna, the rule was reversed with respect to left and right.

Initially, each horse was turned loose before approaching the setting, but Elna paid little attention to the object after a small number of sessions. For this reason, the handler started guiding Elna to the object and prevented her from reaching for the doors before she touched the object or moved her muzzle very close to it. After this procedure was introduced, matching objects were placed on the ground in front of the doors. This experimental configuration resembles the original matching-to-sample task, but in contrast to the original task, the positions of the test stimuli remained unchanged between trials because of the object-to-position rule.

For the altered experimental setting for Elna, inadvertent cueing from the handler cannot be ruled out and apparent learning of the task has to be confirmed with a different experimental procedure. As the time limit for the experiments approached, Elna's task was altered for this reason. For Bruno, the experiment remained unchanged, meaning that only one object (the sample) was placed on the ground and he was turned loose 5 m in front of the setting.

2.6 A simplified matching-to-sample task

As the end of the time frame for the experiments was approaching, the matching-to-sample task was simplified for Diamond and Monica because their performance did not raise significantly above chance level. In the simplified setting, only the symbols ○ and ✕ were used and the position of the test stimuli was not altered between trials. This means that only 2 of the 12 original stimulus configurations were used for each horse and that the task can be solved by symbol-to-position matching. The test stimuli for Diamond were ✕ to the left and ○ to the right, whereas their position were reversed for Monica.

2.7 Procedures to counteract side preferences

In phase one of the initial MTS-training, side preferences were counteracted by the correction trials after each failed trial. In the later phases and in the later experiments, correction trials were applied to a very limited extent. However, a new procedure to counteract a strong side preference was introduced in the experiments following the initial MTS-training. This procedure was applied after selected sessions and the stimuli setup with the correct response on the nonpreferred side was used consistently. Trials were repeated until the horse achieved two consecutive correct responses.

2.8 Statistics

For each trial, the first response of the horse was registered. (No outcomes of correction trials were included in the statistics.) Pushing a door was counted as a selection that was either correct or incorrect. If the horse tried to reach around the plywood wall or walked away, the outcome was recorded as no selection. Missing data correspond to canceled trials, errors in the stimuli setup, and errors in the record keeping.

Binomial tests can be used to quantify the statistical significance of the results. To enable detection of partial learning, I used the following procedure. (I) Let S denote the total number of sessions. For each number $s \in \{1, \ldots, S\}$, consider the s last sessions and calculate the P-value for obtaining at least the actual number of correct selections among the total number of selections. (II) Define P_0 as the minimal P-value obtained in step I. (III) Calculate the probability P to obtain a smaller value of P_0 under the assumption that each selection is correct with probability 0.5. The probability P in the last step is reported as the P-value for performance above chance.

Let n_s denote the total number of selections during the last s sessions and let k_s denote the corresponding number of incorrect selections. Then,

$$P_0 = \min_{s\in\{1,\ldots,S\}} 2^{-n_s} \sum_{m=0}^{k_s} \binom{n_s}{m} . \tag{1}$$

Consider a stochastic process where the variables n and k are initially set to 0 and each step in the process

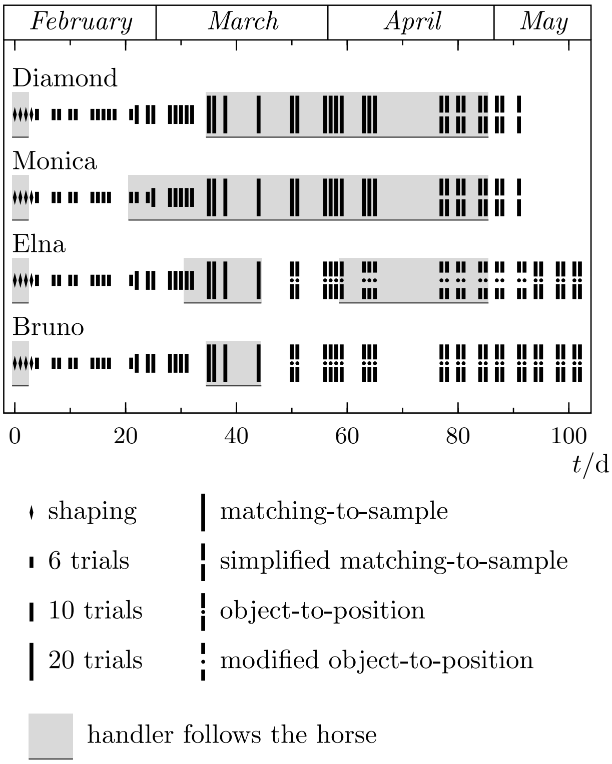

Figure 2. Overview of experiment sessions. The time axis displays the elapsed time t in days relative to February 4 which was the first day of shaping. Modified object-to-position refers to the object-to-position task with the addition of matching objects.

t/d	i_{D}	i_{M}	i_{E}	i_{B}	Event
0	1	1	1	1	start of shaping
3	4	4	4	4	horses turned loose
4	5	5	5	5	MTS-training, phase 1
21	–	14	–	–	handler starts following
22	16	–	15	15	MTS-training, phase 2
25	–	17	–	–	MTS-training, phase 2
31	–	–	21	–	handler starts following
35	24	–	–	22	handler starts following
35	24	23	23	22	MTS-training, phase 3
50	–	–	27	26	object-to-position matching
50	–	–	27	26	horses turned loose
59	–	–	32	–	handler starts following
63	–	–	33	–	addition of matching objects
77	37	36	–	–	simplified MTS
77	37	–	36	–	side-preference compensation after sessions
87	43	42	42	–	horses turned loose
87	43	42	42	–	additional responses encouraged after errors
91	45	44	–	–	end of experiments
94	–	–	46	–	removal of matching objects
102	–	–	51	50	end of experiments

Table 1. Schedule for the experiments. Each event in the schedule represents the onset of a new phase in the experiments or a change in the experimental procedures. The time t, in days elapsed since the first day of shaping, is shown in column 1. Columns 2–5 display session indices i_{D}, i_{M}, i_{E}, and i_{B} of training sessions for Diamond, Monica, Elna, and Bruno, respectively. Each event applies to horses with session indices displayed in the row of the event. The session index is a number that is set to 1 for the first session and increases by one for each session the horse takes part in.

increases n by 1. Let k be increased by 1 or unaltered with equal probabilities for each step. The process stops if there is an $s \in \{1, \ldots, S\}$ such that $n = n_s$ and

$$P_0 \geq 2^{-n} \sum_{m=0}^{k} \binom{n}{m} . \tag{2}$$

Then, P is given by the probability that the process stops. Let $p_{n,k}$ denote the probability that the stochastic process visits the state determined by n and k. Furthermore, let $a_{n,k}$ be 1 if the stopping criterion is satisfied and 0 otherwise. Then,

$$P = \sum_{n,k} a_{n,k}\, p_{n,k} , \tag{3}$$

where the probabilities $p_{n,k}$ are calculated from the recurrence relation

$$p_{n,k} = \frac{1 - a_{n-1,k-1}}{2} p_{n-1,k-1} + \frac{1 - a_{n-1,k}}{2} p_{n-1,k} \tag{4}$$

and the conditions $p_{0,0} = 1$ and $p_{n,-1} = p_{n,n+1} = 0$ for $n = 0, 1, 2, \ldots$.

3 RESULTS

The time line for the experiments is presented by figure 2 and table 1. The horses learned quickly to push a door open and retrieve the reward (shaping), but no learning was evident from the performance in the initial MTS-training, see figure 3. ($P > 0.1$ with the definition in section 2.8.) In the beginning of the matching-to-sample training, the horses spontaneously directed their attention to the sample stimulus but they appeared to lose this attention as the experiment progressed. When it became apparent that the horses did not maintain attention to the sample stimulus, the handler started to follow the horses along while guiding them towards the sample stimulus. The handler also corrected their head position when it appeared to be too low for looking at the sample stimulus. For Elna, this change was introduced on day 31 (i.e., 31 days after the first day of habituation/shaping), meaning that there were 15 sessions of MTS-training before this change. Similar guiding was introduced earlier (day 21) for Monica and later (day 35) for Bruno and Diamond.

Because of the failure of learning the initial task, two horses were selected to perform object-to-position matching tasks with physical objects as stimuli. Elna and Bruno displayed the strongest tendencies to move

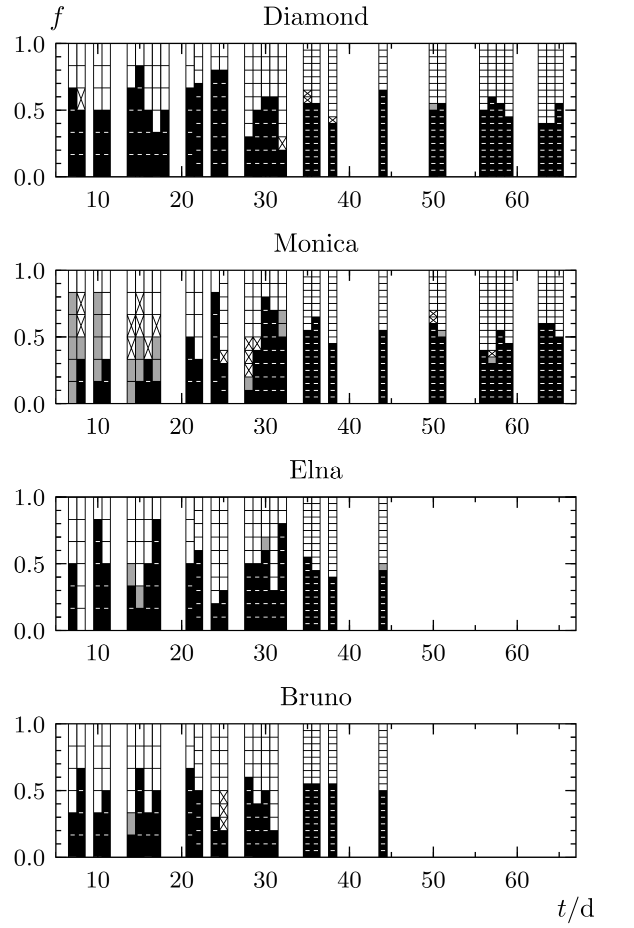

Figure 3. Performance in the initial matching-to-sample training. The time axis displays the elapsed time t in days relative to the first day of shaping. The fraction f of correct responses for each session is displayed in black bars. Empty boxes show the number of incorrect responses whereas crossed boxes and gray boxes, respectively, represent events of no selection and missing data. Missing data are caused by canceled trials and experimental errors.

their attention away from the symbolic stimuli and were hence selected for the altered experiments. The objects in the object-to-position matching were placed on the ground because Bruno spontaneously directed his attention towards things on the ground. The altered experiments started on day 50, and this change limited the number of sessions of MTS-training to 22 and 21 for Elna and Bruno, respectively.

Bruno learned the object-to-position task ($P = 2.4 \cdot 10^{-18}$) but Elna did not ($P > 0.1$). See figure 4. During Bruno's last 10 sessions, he made 154 correct selections and 39 incorrect selections. The remaining 7 outcomes were events of no selection. Both Elna and Bruno paid attention to the stimuli objects initially but as the experiments progressed Elna seemingly stopped to pay attention to them. Bruno touched the sample stimulus quite frequently and developed a habit of biting and dragging the cone along the ground.

As Elna gradually paid less attention to the sample stimulus, the handler started following along and guiding her towards the stimulus. Before each response, she was required to touch or get very close to the sample. At this stage of the experiment, additional objects were placed on the ground in front of each door. This alteration meant that Elna's task could be solved by matching the sample to a test stimulus and then selecting the door above this stimulus. After 9 sessions with the additional objects, the handler turned Elna loose again. With this change, the additional objects seemed to move Elna's attention away from the sample stimulus and they were removed after 4 additional sessions. During the remaining 6 sessions, Elna apparently paid more attention to the sample stimulus but no learning of the task was evident from these trials.

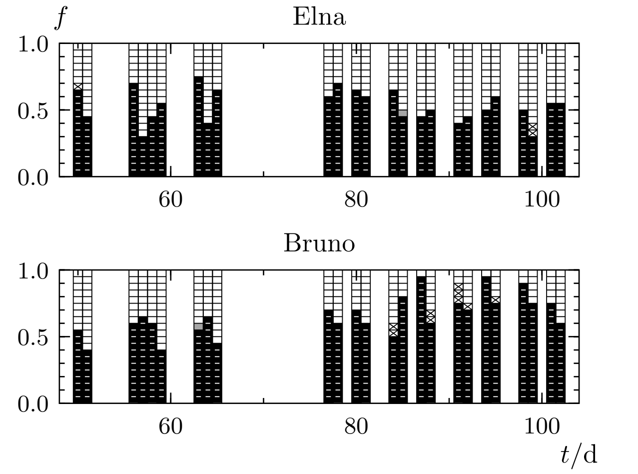

Figure 4. Performance in object-based conditional discrimination training where a cone or a foam block was used as a sample stimulus. Responding to the left or to the right was rewarded depending on the sample. The time axis displays the elapsed time t in days relative to the first day of shaping. The fraction f of correct responses for each session is displayed in black bars. Empty boxes show the number of incorrect responses whereas crossed boxes and gray boxes, respectively, represent events of no selection and missing data.

From the sessions before day 77, there were no indications of performance above chance level for Diamond and Monica in the original MTS-task. For this reason, their task was simplified as described in section 2.6. In the simplified setting, Diamond appeared to increase her attention towards the stimuli and this increase of attention seemed to be easily compromised by the actions of the handler. Because of this problem and similar problems for Monica and Elna, all horses were turned loose on day 87 and later sessions. Note that Bruno already was turned loose in the trials since day 50.

On day 87, it was unclear to what extent the horses had understood that there was always exactly one correct alternative. For this reason, Diamond, Monica, and Elna were allowed and encouraged to make additional responses after each failed trial. Their initial tendency to push the incorrect door repeatedly indicated that the choice situation was not transparent to

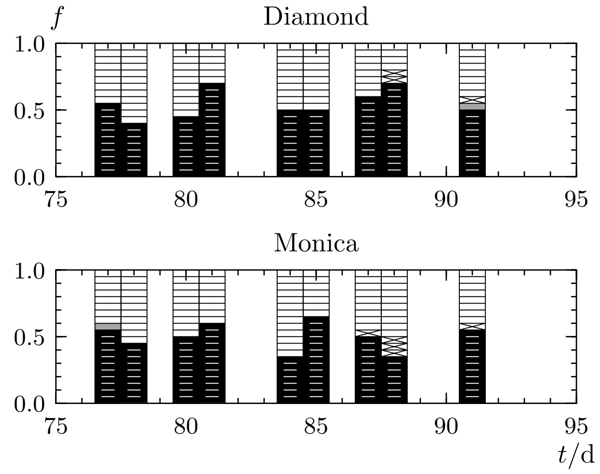

Figure 5. Performance in symbol based conditional discrimination training where two configurations from the initial matching-to-sample training were used for each horse. The configuration of test stimuli remained unchanged between trials for each horse whereas the sample stimulus varied randomly. The time axis displays the elapsed time t in days relative to the first day of shaping. The fraction f of correct responses for each session is displayed in black bars. Empty boxes show the number of incorrect responses whereas crossed boxes and gray boxes, respectively, represent events of no selection and missing data.

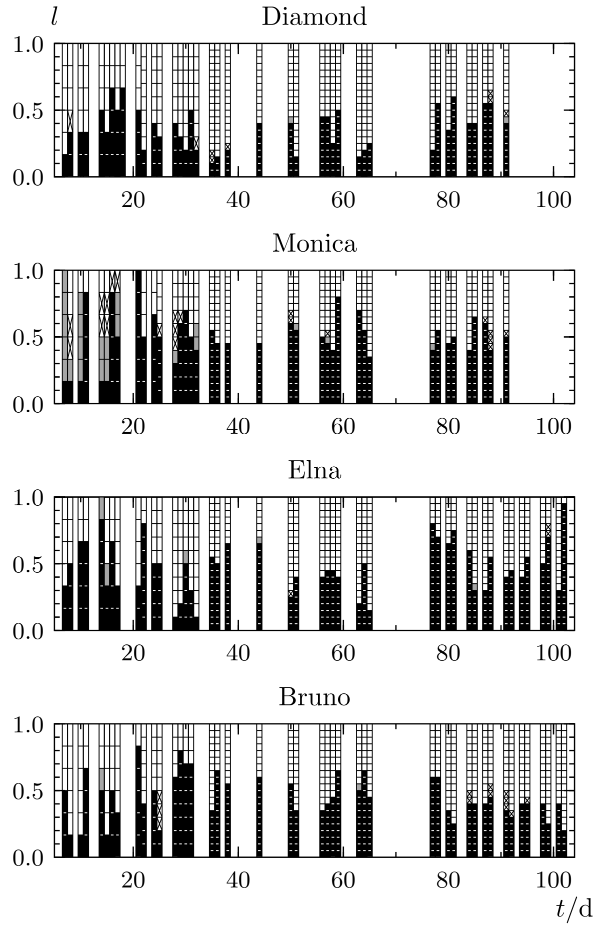

Figure 6. Left-side bias for all experiments. The time axis displays the elapsed time t in days relative to the first day of shaping. The fraction l of responses to the left side for each session is displayed in black bars. Empty boxes show the number responses to the right side whereas crossed boxes and gray boxes, respectively, represent events of no selection and missing data. Missing data are caused by canceled trials and experimental errors.

these horses. Bruno's performance appeared to be significantly above chance at the same time and, hence, the experimental procedure for Bruno was not altered.

For the simplified MTS-task, Diamond's and Monica's performances were not significantly above chance. ($P = 0.06$ for Diamond and $P > 0.1$ for Monica.) See figure 5. There were 32 and 31 sessions, respectively, for Diamond and Monica in the initial MTS-task. After 8 sessions for each horse, the simplified MTS-experiment was stopped, because these horses were not motivated enough for continuing the experiments.

Figure 6 displays the fraction of responses to the left for each horse in all types of trial sessions. Procedures to correct for preference of one side (as described in section 2.7) were applied to Diamond and Elna after their respective trial sessions on day 77. The numbers of correction trials to the stopping criterion (two consecutive correct responses) were 14 and 15 for Diamond and Elna, respectively.

4 DISCUSSION

In the MTS-training, the sample stimulus is likely to catch the attention of the horses for two reasons. First, it is displayed in a more prominent position relative to the test stimuli. Second, the fact that the stimulus is altered from trial to trial is likely to attract some attention. However, there is no mechanism to *maintain* the the horses' attention to the sample stimulus during subsequent sessions. Hence, the horses might learn to ignore the sample stimulus unless they relatively quickly learn that there is some relation between the sample stimulus and the rewarding test stimulus. This means that the failure of the horses to learn the initial matching-to-sample task is not necessarily an indication that horses cannot handle the concept of visual identity: it is likely that the learning of the task was restricted by lack of attention to the sample stimulus.

As pointed out in the introduction, successful MTS-learning is often dependent on a procedure where the subject performs some action directed to the sample stimulus. Activity directed towards the sample stimulus can increase attention and provide a context that promotes learning. For example, Wright and Delius (1994) demonstrated drastic improvements in learning rates for pigeons that dug into colored gravel in comparison to settings where only pecking at the stimuli was required. It is also useful to consider experiments where a procedure for promoting attention has failed. A particularly interesting example is the experiment presented by Persson (2008, chapter 12), where gorillas failed to learn an apparently simple matching-to-sample task. In this experiment, cards with test stimuli were presented after the experimenter had observed that the tested gorilla paid attention to a card with a sample stimulus. This procedure ensured some attention towards the sample card, but it did not ensure that attention was directed towards the *symbolic stimulus* on the card. Furthermore, Persson (2008, chapter 12) presents variations of the experiment that indicate that the gorillas ignored the visual information printed on the cards.

The object-to-sample matching task was inspired by Bruno's interest in objects on the ground. The fact that Bruno spontaneously directed his actions towards the sample stimulus in the altered setting seems to be crucial to his learning of the task. In particular, he developed a habit of biting and dragging the cone along the ground and he apparently associated these actions with selection of the left door (which was the correct selection for the cone stimulus).

In the object-to-position matching experiments, Bruno responded correctly 16 times out of 20 on day 85. His best performance before this session was 14 correct responses out of 20. Bruno's behavior before day 85 indicated partial learning of the task. During days 77–78, his response to the cone and the block was correct in 13 of 16 and 13 of 24 trials, respectively. These results indicate that he had begun to associate the cone with the correct response without developing a side preference that decreased the fraction of correct responses to the block. During the following two sessions, the trend was reversed for the stimuli with 9 of 20 and 17 of 20 correct answer for the cone and the block, respectively.

One interesting aspect of the statistics is to what extent such partial learning is evident from the data collected from the 14 sessions before day 85. For such a comparison, the P-values for the above mentioned events are very likely to be misleading, because they were selected after inspection of the results. However, we can calculate the P-value according to the definition in section 2.8 in the case that day 84 were the last day of the experiment. Such a calculation gives $P = 0.014$ which, on the one hand, is below the standard P-value threshold of 0.05. On the other hand, if day 84 really were the last day of the experiments, the corresponding P-value for any of the four horses to obtain such a result by chance is slightly above 0.05.

Taking Bruno's learning performance into consideration, it is plausible that Diamond or Monica could have learned the simplified MTS-task if they were given enough time under suitable circumstances. For these horses, the rewards used in the experiments were apparently too weak in comparison to the fresh grass in their pasture and they were not motivated enough for continuing the experiments. For this reason, the best time for conducting this type of experiments is probably during the autumn and winter. In this project, the experiments started in February and ended in May.

Because the horses failed to learn the initial task and the time frame for the project was too narrow to systematically investigate learning of related tasks, the immediate importance of the results is limited. However, interesting trends were observed and the failure of learning the initial task indicates that there is a need for more elaborate procedures for training horses in matching-to-sample tasks. In future research, it would be interesting to see under what circumstances horses learn a task where a sample is matched to a position. For example, is such learning facilitated or obstructed by having additional stimuli displayed in such a way that the task can be solved by IMTS learning? Also, the correction procedures are likely to play a crucial role in the learning. Allowing additional responses after an incorrect response might help learning, but it might also decrease the incentive to select the correct alternative at the first response in each trial.

In the main part of this project, a trial was either ended with no additional attempts or with correction trials in the form of additional trials for the same stimulus configuration. Tests carried out on day 87 showed that Diamond, Monica, and Elna tended to try the same response repeatedly when they were left standing close to the equipment or were backed a short distance away from it. This behavior gives a clear indication that these horses had not learned that exactly one door is unlocked during each trial. Hence, there is a need to investigate what horses can learn from various situations where they are allowed to make additional responses after an initial incorrect response. Basic understanding of such correction procedures and their effects is likely to be crucial for future research about learning behaviors in horses. Furthermore, understanding of correction procedures, in controlled settings, might yield insights that are important to horse training in general.

5 SUMMARY

In this project, I tested learning of an identity matching-to-sample task for four horses. Because none of the horses succeeded in the initial task, the extent of identity-related concept formation could not be investigated. One horse, Bruno, did succeed in learning of a simpler task where each one of two physical object were to be matched to a response to the left side or to the right side of the experimental setting.

The amount of attention to the sample stimulus was apparently a crucial component in the failure of learning the initial task and in Bruno's successful learning of the object-to-position matching task. All horses did initially pay attention to the sample stimulus in the first task but there was no mechanism for the horses to maintain their attention and, most likely, they learned to ignore the sample stimulus. In Bruno's second task, attention to the sample stimulus was overall maintained because he did often touch the sample. Furthermore, Bruno developed a habit of biting one sample object (the cone) and this habit was probably important to his learning of the task.

Although the project failed with respect to its original purpose of testing identity-related concept formation in horses, the results can be valuable for future research on this topic. In particular, the experimental setting was similar to settings where horses successfully had learned categorization tasks and to select one of two stimuli by comparing sizes or hues of the stimuli. It is generally a good idea to keep the experimental procedure as simple as possible, but the negative result for the identity matching-to-sample task motivates testing of more elaborate procedures when a sample stimulus is introduced.

ACKNOWLEDGMENTS

I thank the department of Cell and Organism Biology for enabling this project, with particular thanks to Almut Kelber and Lina Roth for help with the experiments and the experimental design. I thank Christer, Kicki, and Marcus Odell at Esarps Quarter & Paint for letting me borrow their horses and for providing a very good environment for the experiments. I also thank Ture and Cecilia Samuelsson, Johan Jernström, and Frida Bruhn for help with the experiments.

D'Amato, M. R., Salmon, D. P., and Colombo, M. (1985), *Extent and limits of the matching concept in monkeys (Cebus apella)*, Journal of Experimental Psychology: Animal Behavior Processes **11**, 35–51.

Flannery, B. (1997), *Relational discrimination learning in horses*, Applied Animal Behaviour Science **54**, 267–280.

Geisbauer, G., Griebel, U., Schmid, A., and Timney, B. (2004), *Brightness discrimination and neutral point testing in the horse*, Canadian Journal of Zoology **82**, 660–670.

Hanggi, E. B. (1999), *Categorization learning in horses (Equus caballus)*, Journal of Comparative Psychology **113**, 243–252.

Hanggi, E. B. (2003), *Discrimination learning based on relative size concepts in horses (Equus caballus)*, Applied Animal Behaviour Science **83**, 201–213.

Katz, J. S., Wright, A. A., and Bachevalier, J. (2002), *Mechanisms of same/different abstract-concept learning by rhesus monkeys (Macaca mulatta)*, Journal of Experimental Psychology: Animal Behavior Processes **28**, 358–368.

Lombardi, C. M. (2008), *Matching and oddity relational learning by pigeons (Columba livia): transfer from color to shape*, Animal Cognition **11**, 67–74.

Macuda, T. and Timney, B. (1999), *Luminance and chromatic discrimination in the horse (Equus caballus)*, Behavioural Processes **44**, 301–307.

Mauck, B. and Dehnhardt, G. (2005), *Identity concept formation during visual multiple-choice matching in a harbor seal (Phoca vitulina)*, Learning & Behavior **33**, 428–436.

Murphy, J. and Arkins, S. (2007), *Equine learning behaviour*, Behavioural Processes **76**, 1–13.

Oden, D. L., Thompson, R. K. R., and Premack, D. (1988), *Spontaneous transfer of matching by infant chimpanzees (Pan troglodytes)*, Journal of Experimental Psychology: Animal Behavior Processes **14**, 140–145.

Peña, T., Pitts, R. C., and Galizio, M. (2006), *Identity matching-to-sample with olfactory stimuli in rats*, Journal of the Experimental Analysis of Behavior **85**, 203–221.

Persson, T. (2008), *Pictorial primates: a search for iconic abilities in great apes*, Ph.D. thesis, Lund University, Department of Cognitive Science.

Roitberg, E. and Franz, H. (2004), *Oddity learning by African dwarf goats (Capra hircus)*, Animal Cognition **7**, 61–67.

Roth, L. S. V., Balkenius, A., and Kelber, A. (2007), *Colour perception in a dichromat*, Journal of Experimental Biology **210**, 2795–2800.

Vonk, J. (2003), *Gorilla (Gorilla gorilla gorilla) and orangutan (Pongo abelii) understanding of first- and second-order relations*, Animal Cognition **6**, 77–86.

Wright, A. A. (2001), *Learning strategies in matching-to-sample*, in Cook, R. G. (editor), *Avian visual cognition*, [On-line] www.pigeon.psy.tufts.edu/avc/wright/.

Wright, A. A. and Delius, J. D. (1994), *Scratch and match: pigeons learn matching and oddity with gravel stimuli*, Journal of Experimental Psychology: Animal Behavior Processes **20**, 108–112.

Wright, A. A. and Katz, J. S. (2006), *Mechanisms of same/different concept learning in primates and avians*, Behavioural Processes **72**, 234–254.