

Long-term memory for categories and concepts in horses (*Equus caballus*)

Evelyn B. Hanggi · Jerry F. Ingersoll

Received: 7 September 2008 / Revised: 5 November 2008 / Accepted: 14 December 2008 / Published online: 16 January 2009
© Springer-Verlag 2009

Abstract Three horses (*Equus caballus*) with a history of performing cognitive tasks including discrimination learning, categorization, and concept use were tested to evaluate their long-term memory (LTM) in three experiments. In addition, use of LCD multi-displays for stimulus presentation was incorporated into cognition testing protocol for the first time with horses. Experiment 1 tested LTM for discrimination learning that originally occurred 6 years earlier. Five sets of stimuli were used and the two horses tested showed no decrement in performance on four of the sets; however, both horses did score below chance on one set. Experiment 2 examined long-term categorization recall 10 years after horses had demonstrated the ability to make stimulus selections based on shared characteristics within a given category. The horse tested for LTM after the decade-long interval immediately and consistently applied the previously learned categorization rule to not only familiar but also novel sets of stimuli. Experiment 3 tested another horse for LTM for a relative size concept. This horse had originally demonstrated concept rule use in order to select stimuli based on their relative size to one another. More than 7 years later and without further training, this horse reliably applied the previously established size concept to both familiar and novel sets of stimuli. These findings are the first reports of long-term categorical and conceptual memory in horses and are consistent with observations of domestic and wild horses, which indicate that behavioral and ecological events may be remembered for long periods of time. These studies also demonstrate the adaptive nature

of horses with regard to their ability to generalize over several different testing conditions.

Keywords Horse · Long-term memory · Concept · Categorization · Discrimination · LCD display · Equine

Introduction

Intrinsic to the success of almost any cognitive operation is the role of memory. Without short- and long-term memory (LTM), higher-level problem solving (not to mention most basic learning) in animals would be impossible. Cognition, learning and memory factor into almost every aspect of daily life for domestic horses (*Equus caballus*) as well as for feral or wild horses and other equids. For their survival and success, free-ranging horses must learn and remember their social, biological and physical environments, which include conditions that predictably change at times and not at other times (Linklater 2007), in addition to being capable of adaptive modification of their innate abilities. Conspecific interactions and social relationships, foraging, predation, homing, and climatic variation make it necessary for individuals to possess and utilize at least some degree of cognitive complexity. Domestic horses may not face as severe survival challenges as their wild counterparts but are often confronted with handling and training practices that challenge their mental and physical well-being. On a more benign level, domestic horses must still learn about, understand and remember both inter- and intraspecific relations, schedules, and rules based on natural or artificial events. Horses with the greatest capacity to learn, understand, and solve problems are more likely to succeed with regard to the human–horse relationship and the associated handling and training atmosphere (Murphy and Arkins 2007).

E. B. Hanggi (✉) · J. F. Ingersoll
Equine Research Foundation,
P.O. Box 1900, Aptos, CA 95001, USA
e-mail: EquiResF@aol.com
URL: <http://www.equinersearch.org>

Given the anecdotal reports, it appears that the horse possesses an excellent memory and recall ability. The training of horses provides ample opportunity to observe how well horses remember repetitive events and an individual's anxious reaction to situations that years earlier had frightened it suggests a remarkable LTM (Waring 2003). The evidence that has been reported from the scientific community supports these beliefs. Horses have demonstrated the ability to remember correct choices in multiple two-choice discrimination tasks for several months (Giebel 1958; Dixon 1970) as well as recall a learned response under maze conditions after a period of 1 week (Marinier and Alexander 1994). With respect to short-term memory, Grzimek (1949) found that one horse could correctly solve a delayed response test for food location after 6 s, while another was able to delay up to 60 s. Using a two-choice delayed-reaction procedure, Nobbe (1978) showed that a filly could achieve a 24-s delay with high accuracy. Most recently, and contrary to a report by McLean (2004), Hanggi (unpublished data) found that, under appropriate testing circumstances, horses could solve spatial short-term memory tasks for food targets and abstract stimuli following delays of 5, 10, 15, 20 and 30 s with great accuracy. The disparate outcomes of these last two studies appeared to be the result of differing methodologies, subjects, and training history.

Despite these studies, the limits of equine memory—especially recall involving higher-order cognitive skills—have not been fully investigated. Indeed, for nonhuman animals, information on categorical and conceptual LTM is very limited. This is primarily due to the longitudinal requirements of such studies, originating with the meticulous collection of training and testing data and ending with retesting several months to many years later. Because of the progressive nature of animal cognition research, as well as the complexity of maintaining specific testing conditions and test subject availability over extended periods of time, LTM testing in nonhuman animals occurs only rarely. Therefore, when opportunities arise to obtain and report LTM data investigators are urged to do so (Burdyn et al. 1984).

The handful of studies that have explored LTM in subjects with conceptual learning experience have been done primarily with primates. Eight rhesus monkeys (*Macaca mulatta*) performed nearly perfectly on oddity learning sets 7 years after original testing (Johnson and Davis 1973). Three squirrel monkeys (*Samiri sciureus*) demonstrated retention of a relational concept after more than 2 years and one other showed exceptional retention after 5 years (Burdyn et al. 1984). Four gorillas (*Gorilla g. gorilla*) performed as well on retesting 2.5 years after initial testing on discrimination reversal problems (Patterson and Tzeng 1979). Until now, the only nonprimate conceptual LTM study was done using one California sea lion (*Zalophus californianus*). This sea lion showed no decrement in retention of an

associative concept 1 year after testing and immediately and repeatedly demonstrated the use of an identity concept to familiar and novel sets of stimuli in a 10-year memory test (Reichmuth Kastak and Schusterman 2002).

As research into equine LTM for higher-order cognitive functions was nonexistent, and because we had just modified our testing apparatus to include LCD multi-displays, we took this opportunity to test our horses' LTM for stimulus discrimination, categorization, and concept usage.

Methods

Animals

Three horses were tested for LTM in three different experiments. Bodie, a 13-year-old Pinto Draft mix gelding, participated in a discrimination learning memory test; Coco Bean, a 13-year-old Paint gelding, participated in the discrimination learning memory test as well as a categorization memory test; and Tequila, a 13-year-old Arabian gelding participated in a size concept memory test. These horses lived with nine others at the 16-ha Equine Research Foundation and participated in unrelated experiments in the interim between the original experiments and the LTM tests. Bodie came to the Foundation as a 6-year-old and Coco Bean and Tequila were donated as 1- and 2-year-olds, respectively. They were fed 7.2 kg grass hay daily and were not food deprived. When not involved in research, these horses were either in pasture with herd mates, being trained or ridden.

Experiment 1: LTM for discrimination learning

Original procedure

The first LTM test derived from individual two-choice discrimination tasks using five sets of stimuli that had been learned during a picture/object recognition experiment more than 6 years earlier (Hanggi 2001). Bodie and Coco Bean had originally learned to discriminate between three-dimensional (3D) objects and also between two-dimensional (2D) photographs of these objects. The 3D stimuli were children's toys or household objects of different shapes and colors: red plastic Plate; red foam Pillar; green and yellow plastic measuring Scale; purple, yellow, white, and red dog-shaped Frame; green and orange rubber toy Hedgehog; blue, green, and red plastic Bottle; multi-colored cloth covered foam Football; multi-colored cloth covered foam Frisbee; turquoise Tyco Cookie Monster (CM) toy; brown Tyco Snuffleupagus (Gus) toy. The 2D stimuli were laminated photographs of these objects taken with a Nikon Coolpix 600 digital camera—with extraneous information removed using Adobe Photoshop—and printed

on a Hewlett-Packard 842C Color printer. Each 2D stimulus was placed into a 33-cm square Plexiglas folder transparent on the viewing side and white on the back.

The apparatus consisted of a 1.9×2.4 m (height \times width) wooden wall containing five 30-cm square openings and was located in the breezeway of a well-lighted stable. Each opening held an opaque plastic panel that could slide horizontally to reveal the 2D stimuli and each had a shelf located below it on which the 3D objects could be placed. Two hidden assistants controlled placement of the stimuli and panel movement from behind the wall. A station, located 3.6 m in front of the apparatus, served as a holding area where the horse stood unattended until it was given a release cue to approach the apparatus. This release was the lowering of a horizontal bar that spanned in front of the horse and was controlled from behind the apparatus. Upon release, the horse walked by itself up to the apparatus and selected one of the two stimuli of a given set by touching it with its nose. If correct, food reinforcement (15 g of mixed grain) was delivered from behind the apparatus via a chute to a feed bowl centered at the base of the apparatus. If incorrect, the horse was told “No” and was not given a food reinforcer. After eating the grain or hearing “No,” the horse returned to station on its own to await the next trial.

During this experiment, which was conducted over a 3-month period, the horses were presented with sessions consisting of approximately 30 trials that lasted about 40 min. Both horses initially performed at chance levels for all training sets except on the first set, for which the preferred stimulus was designated as the correct choice. All subsequent positive stimuli were nonpreferred in preference tests. The horses then learned to discriminate between stimuli through trial-and-error within 20–80 trials. Once discrimination was acquired (19/20 correct consecutive responses; $z = 4.02$, $P = 0.00002$, $\alpha = 0.01$, binomial test) and additional overlearning trials were run, the corresponding 2D or 3D replica was tested. This was followed by the training and testing of another set of stimuli. For some stimulus sets, the 3D discrimination was learned first; for others, the 2D discrimination came first. After acquisition, both horses transferred their responses between dimensions for most of the sets, i.e., they immediately and consistently selected the photograph of the correct 3D object or vice versa (Hanggi 2001). Bodie and Coco Bean’s last testing of these stimuli occurred in May 2001.

Procedure for testing horses using LCD multi-displays and a Macintosh computer

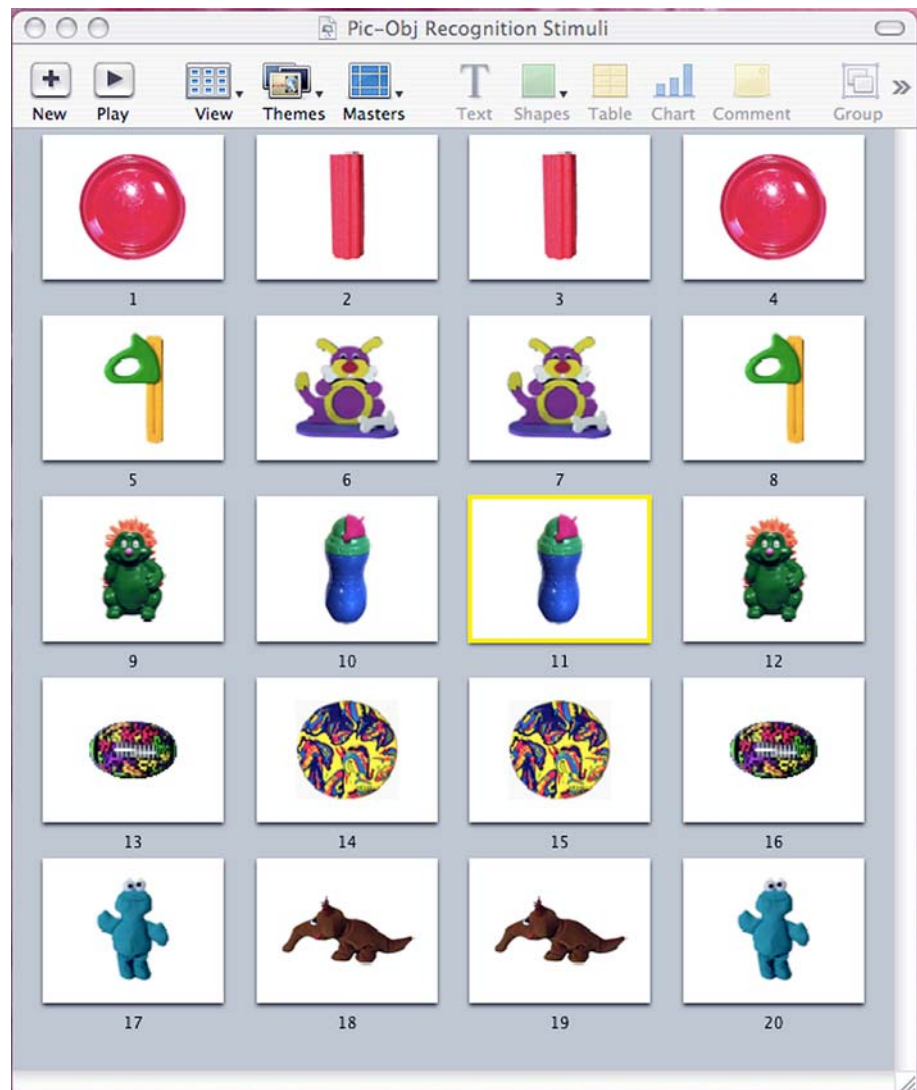
Prior to this study, the testing apparatus described above was the standard system at the Equine Research Foundation (Hanggi 1999a, b, 2001, 2003, 2006, 2007; Hanggi et al. 2007). In 2007, we designed and constructed a new apparatus

that contained LCD computer screens whereupon stimuli were directly displayed. To the best of our knowledge, computer displays had never been used before with horses; thus, we did not know how well horses could perceive images on them. The new apparatus comprised a 2.0×2.4 m painted wooden wall, which was placed in the same location in the stable as the old one. It contained two 29.2×36.8 cm openings, placed 107 cm apart and 81.3 cm from the bottom. A 48.3-cm ViewSonic VX922 LCD Display was set into each opening and attached to the back of the apparatus with rubber strapping. Clear nonglare Plexiglas windows were attached over the openings on the front of the apparatus to protect the LCD displays from contact with the horses. The displays were connected by video cables to a Matrox DualHead2Go Multi-Display Upgrade box (Matrox Graphics, Inc.), which was further connected by a 15.2 m VGA cable to a 17" MacBook Pro computer. Display resolution was set to $2,560 \times 1,024$, refresh rate 60 Hz, and computer resolution was set to $1,680 \times 1,050$. The station was the same as in previous experiments.

The application Keynote (Apple Computer, Inc.) was used to organize and present stimuli on the LCD displays. The stimuli consisted of computer images of the color photographs used in the original picture/object recognition study. For the other experiments described in this paper stimuli were computer-generated images. Keynote has the capability of displaying stimuli as thumbnail icons on the MacBook Pro computer screen in a view option similar to a light table (Fig. 1). This allows the experimenter to easily view and select individual stimuli from a large collection of images and present them on either LCD display. Data may be recorded simultaneously on the same computer or separately.

Stationing and response behaviors were originally taught using the Equine Research Training System™ (ERTS; Hanggi 2006, 2007), which enables horses to work unrestrained and unattended, thereby minimizing the risk of cueing by human assistants as well as minimizing the number of personnel (Hanggi 1999a, b, 2001, 2003, 2005, 2006, Hanggi et al. 2007). A trial began when the experimenter, located behind the apparatus, selected the desired stimulus set (e.g., 1 and 2 or 11 and 12; Fig. 1; position set according to a pre-determined random series) on the Keynote window. After a 5-s observation period, the horse was released from the station, walked up to the apparatus and lightly touched the Plexiglas window covering the stimulus on one of the displays. Responses to stimuli were reinforced as in previous experiments and were observed from behind the apparatus via video cameras connected to monitors. At the end of each trial, the LCD displays switched to a solid gray screen. Intertrial intervals were dependent on length of time it took for the horse to consume the food reinforcer and to return to station (generally 40–60 s).

Fig. 1 Stimuli used to test LTM for discrimination learning with Bodie and Coco Bean in Experiment 1. This figure also demonstrates use of the Keynote application on a Macintosh computer. Row 1 depicts Plate and Pillar; row 2 depicts Scale and Frame; Row 3 depicts Hedgehog and Bottle; row 4 depicts Football and Frisbee; row 5 depicts Cookie Monster and Snuffleupagus. Each stimulus set appears twice per row so that the experimenter may select stimulus location on the LCD multi-displays. For example, when the experimenter selects 13 and 14, the football appears on the left display and the Frisbee appears on the right. The reverse occurs when 15 and 16 are selected. This allows for flexibility, especially during early training of animals



Procedure for LTM test

Bodie and Coco Bean's memory tests for this experiment were conducted in July 2007, 6 years and 2 months after the original experiment ended. The same photographs that were used in the 2001 experiment were displayed on the LCD computer screens, so the mode of presentation was the only difference between the current test and the original. The horses' ability to remember which stimuli were designated correct in the original study was tested over two sessions comprising ten trials per stimulus set. Neither horse had been presented with these stimuli after completion of the original experiment.

Results and discussion

Both horses had solved the original discrimination problems as part of a picture/object recognition study in 2001. On most training sets they showed a learning curve typical

for trial-and-error acquisition—starting at or below chance levels and then improving until performance met criterion, which generally occurred within 40 trials. Over 6 years later, without any subsequent exposure to these stimuli, both horses chose the correct stimulus on Trial 1 for each of the sets (Table 1). Coco Bean's performance on four of the five sets was nearly perfect, making no errors for the Plate/Pillar discrimination and only one error per set for the Scale/Frame, Football/Frisbee and CM/Gus discriminations. However, he scored only 50% correct on the Hedgehog/Bottle discrimination. Bodie performed as well as Coco Bean on the Plate/Pillar discrimination. His performance was not quite as good on the Scale/Frame, Football/Frisbee and CM/Gus discriminations but still significantly above chance at $\alpha = 0.05$. His performance on the Bottle/Hedgehog discrimination was even worse (20% correct) than that of Coco Bean's and far below chance. During the original experiment, both horses scored at or above 95% correct after learning this set. Individual differences may

Table 1 Performance by two horses on long-term memory tests for five discriminations learned 6 years earlier

Horse	Stimulus set (S+/S-)	CR	Total	% CR	Trial 1	Binomial (z)
Coco Bean	Plate/Pillar	10	10	100.0	CR	3.162 ^a
	Scale/Frame	9	10	90.0	CR	2.530 ^a
	Hedgehog/Bottle	5	10	50.0	CR	0.000
	Football/Frisbee	9	10	90.0	CR	2.530 ^a
	CM/Gus	9	10	90.0	CR	2.530 ^a
Bodie	Plate/Pillar	10	10	100.0	CR	3.162 ^a
	Scale/Frame	8	10	80.0	CR	1.897 ^b
	Bottle/Hedgehog	2	10	20.0	CR	-1.897
	Football/Frisbee	8	10	80.0	CR	1.897 ^b
	Gus/CM	8	10	80.0	CR	1.897 ^b

CR correct response,
% CR percent correct,
S+ positive stimulus,
S- negative stimulus

^a $\alpha = 0.01$, ^b $\alpha = 0.05$, binomial test

account for Coco Bean's slightly higher scores. It is also worthwhile to note that, as in the original experiment, Bodie was the dominant horse of the herd and did at times during intertrial intervals attend to the other horses that were outside. During the LTM experiment, Bodie also developed an intermittent hoof problem, which affected how he walked at times. Both of these factors may have affected his concentration on the task and, hence, his performance.

It is possible that Bottle and Hedgehog appeared more similar to each other than did the stimuli in the other sets. However, in the original experiment, both horses learned this discrimination to criterion so stimulus confusability was not a factor at that time. Moreover, in the original experiment, both horses had as much exposure to this set as they did to the other sets, so lack of experience with the specific stimuli was not a factor in the performance. In the LTM experiment, stimulus sets were tested in the same order they were learned initially with the Hedgehog/Bottle and Bottle/Hedgehog sets occurring in the middle of the five. It is interesting that both horses failed to remember this discrimination but did remember previous and later discriminations. The horses' performance on these tests may indicate primacy and recency effects wherein events learned first and last are recalled more readily than those learned in between (Vauclair 1996). This correlates with rote memory behavior involving serial position effects seen in other species including black-capped chickadees (Crystal and Shettleworth 1994) and pigeons and baboons (Fagot and Cook 2006).

Learning and memorization of arbitrary lists of stimuli and individual discriminations is considered categorization by rote, which is the first level in a classificatory scheme of categorical behaviors (Zayan and Vauclair 1998). Economical principles, including primacy and/or recency effects, may facilitate the memorization of large amounts of information; otherwise, the attentional demands and memory storage capabilities would need to be enormous. Whether the findings from Experiment 1 actually do show serial

position effects in horses remains uncertain. The primary goal of this experiment was to determine whether or not the horses remembered any discrimination at all after 6 years. As is sometimes the nature of opportunistic LTM testing, researchers must make do with what is available. The number of discriminations tested for LTM in this study was only five because that was what was learned as part of the original experiment. This is not enough to make unequivocal conclusions regarding serial position recall. Nevertheless, the fact that two horses behaved so similarly to the learning order warrants further investigation into this aspect of memory.

These findings do indicate that horses can remember events and details over a considerable period of time. Previous research showed that horses could remember discriminations for a number of months but this study demonstrates that recall of this type lasts for years. This must be taken seriously by the horse community when training and handling horses. When working with an animal that learns and remembers as well as the horse does, it is imperative to train it right the first time. The law of primacy does indeed appear to play a critical role in training and management because 'first learned is best learned' (Murphy and Arkins 2007). When what is 'first learned' is a positive experience subsequent learning as well as human/horse interactions are facilitated.

Experiment 2: LTM for categorization

Discrimination problem solving involves learning about explicit relationships that exist between arbitrary stimuli and responses. Each stimulus set must be examined, understood and the correct choice discovered through trial-and-error. Sufficient exposure to multiple stimuli and problems allows animals to improve performance on subsequent problems. As they learn how to solve individual problems they also gain knowledge about the problems' general nature and "solvability." In other words, they learn to learn.

Table 2 Coco Bean's performance on long-term memory tests for categorization using familiar and novel stimulus sets

Familiar sets (S+/S−)	CR	Total	% CR	Trial 1	Binomial (z)
Open square/solid	30	30	100.0	CR	5.477 ^a
Open hexagon/solid	30	30	100.0	CR	5.477 ^a
Open circle/solid	30	30	100.0	CR	5.477 ^a
Open triangle/solid	30	30	100.0	CR	5.477 ^a
Novel sets (S+/S−)	CR	Total	% CR	Trial 1	Binomial (z)
Open hippo/solid	29	30	96.7	IR	5.112 ^a
Open concertina/solid	30	30	100.0	CR	5.477 ^a
Open inkblot/solid	30	30	100.0	CR	5.477 ^a
Open stardrop/solid	29	30	96.7	CR	5.112 ^a

CR correct response, IR incorrect response, % CR percent correct, S+ positive stimulus, S− negative stimulus

^a $\alpha = 0.01$, binomial test

The learning to learn phenomenon (Harlow 1949), wherein animals draw from previous experience to improve later learning, has been observed in several equine studies (Warren and Warren 1962; Fiske and Potter 1979; Sappington and Goldman 1994; Hanggi 1999a, 2003). For example, in concurrent discrimination tests, performance often improved with the addition of subsequent sets, which indicated that the horses understood the experimental requirements. However, first trial response still remained at chance levels, since no rule for problem solving existed. The ability to categorize removes this unpredictability so that errorless responding from the onset is possible.

Filtering out random occurrences with the purpose of recognizing patterns that signal ethological concepts such as food, enemy, or sexual partner is a process of abstraction (Huber 1995). Central to perception and essential for adaptation to changing environments is the capacity to detect invariances that reflect general characteristics of objects and events. Grouping objects or events enable an organism to react to stimuli suitably and efficiently, which is the biological meaning of categorization (Huber 1995). Categories that possess common perceptual properties make up one element of conceptual behavior (Roberts 1998) wherein the application of a concept results in the inclusion of new exemplars that belong to a certain category and the exclusion of those that do not. Categories that fall within perceptual concepts include “trees,” “flowers,” “dogs,” and specialized abstract stimuli as those used in cognitive studies (Hanggi 1999a). Zayan and Vaclair (1998) classify open-ended categories as a level above categorization by rote and note that rules based on principles of perceptual similarity affect sorting behaviors for objects within a given class. Generalization to novel stimuli of the same type is possible through use of the similarity principle.

In Experiment 2, we examined equine LTM for categorization and response to novel stimuli based on a category not seen in a decade (Hanggi 1999a).

Original procedure

Two different apparatuses were used during the original experiment: one was located outdoors in a round pen and another was inside a stable (see Hanggi 1999a for detailed descriptions). Both were dissimilar to the new apparatus used to test LTM. For subsequent unrelated experiments following the one on categorization learning until 2007, the horses primarily used the apparatus described in the original procedure for Experiment 1. Stationing, stimulus presentation, observation and selection, reinforcement, and session length were in accordance to all training and testing done at the Equine Research Foundation (Hanggi 1999a, b, 2001, 2003, 2006, 2007; Hanggi et al. 2007).

The stimuli were computer-generated black images printed on transparent adhesive backed Repro film (made by Rayven, Inc.) and adhered to 33 cm × 37 cm white plastic panels. The correct category consisted of various shapes with an open center; the incorrect category contained all solid shapes. These stimuli were used for two reasons: (1) to determine whether or not horses could categorize objects never before encountered in their environment, which would demonstrate adaptability to a range of situations broader than the class of natural environments (Huber 1995), and (2) in order to avoid introducing irrelevant information as can occur when using pictures of natural scenes.

The findings of the original study indicated that horses are able to sort stimuli categorically. After initial trial-and-error learning during the first discrimination, which took between 80 and 90 trials (learning criterion of 20/20 correct consecutive responses, $z = 4.47$, $P < 0.00001$, $\alpha = 0.01$, binomial test), the horses learned new discriminations rapidly and selected stimuli according to category for 16 stimulus sets. In addition, once the category was understood, they responded accurately on most first trial presentations. They were also able to select the correct stimulus regardless of what it was paired with or how it was oriented. This experiment was last run in 1997.

Procedure for LTM test

In 2007, Coco Bean's LTM was tested using the LCD multi-display apparatus for the first time (Fig. 2a). A subset of the stimuli from the original categorization study was used and, as always, sessions were balanced for location of the correct stimulus choice so that there were no sequence or positional cues available. The first two sessions were comprised trials using square and hexagon as the first two

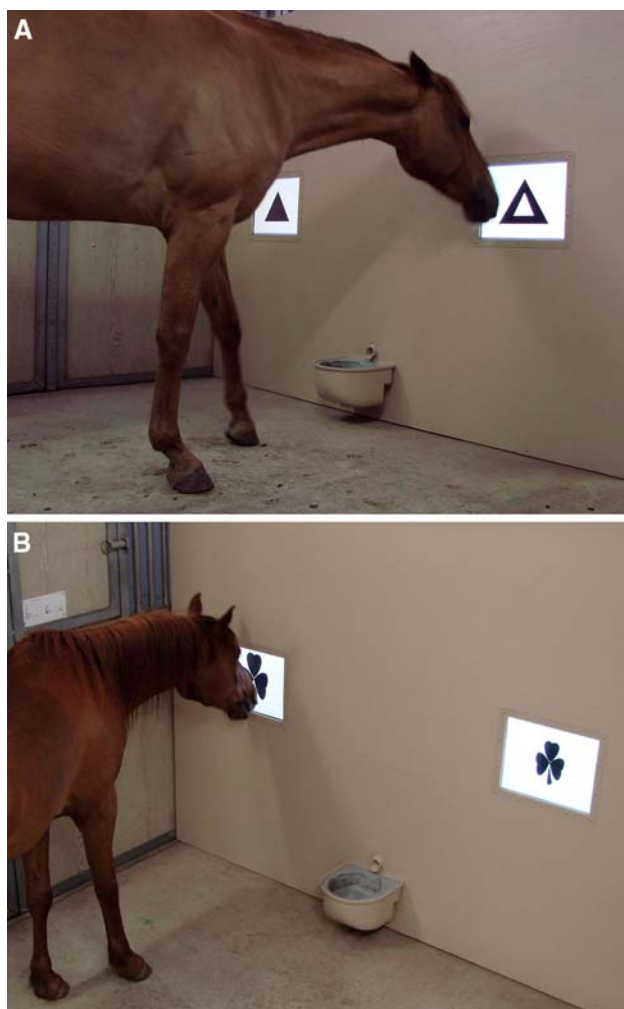


Fig. 2 **a** Coco Bean selecting the correct open-center stimulus during the experiment for LTM for categorization. **b** Tequila demonstrating LTM for a relative size concept on the new apparatus. These experiments were conducted on the new apparatus containing LCD multi-displays

discrimination sets followed by circle and triangle (30 trials per set, first four rows; Fig. 3). Following testing with these four original sets, four novel stimulus sets of more intangible shapes were incorporated into the remaining sessions (30 trials per set, last four rows; Fig. 3). Stimuli were designed using Keynote and balanced for brightness, as were the stimuli in the original categorization experiment.

Results and discussion

In the LTM test, Coco Bean's performance on the old familiar stimuli (square, hexagon, circle, triangle) was perfect (Table 2). This showed that he remembered either each of the individual problems or that he remembered and applied the categorization concept that he had developed in the original experiment. His near perfect overall response to the four novel stimulus sets (hippo, concertina, inkblot, and

stardrop) in addition to correct first trial responses for three out of four sets confirmed that Coco Bean remembered and applied the category established 10 years earlier, rather than relying on rote memory. In fact, the sole Trial 1 error occurred upon initial presentation of the first novel set (hippo) and, thus, may have been due more to the horse suddenly seeing a new unexpected stimulus rather than lack of categorization. This experiment provides strong evidence that horses can remember categories over extended time periods. Coco Bean's performance was nearly perfect prior to cessation of the original experiment and his performance showed no decay 10 years later when testing resumed. These results demonstrate the flexible and durable nature of equine categorical perception and are consistent with our observations of wild and domestic horses that indicate the use of LTM for biologically relevant experiences, individuals, and locations.

Experiment 3: LTM for relative size concept

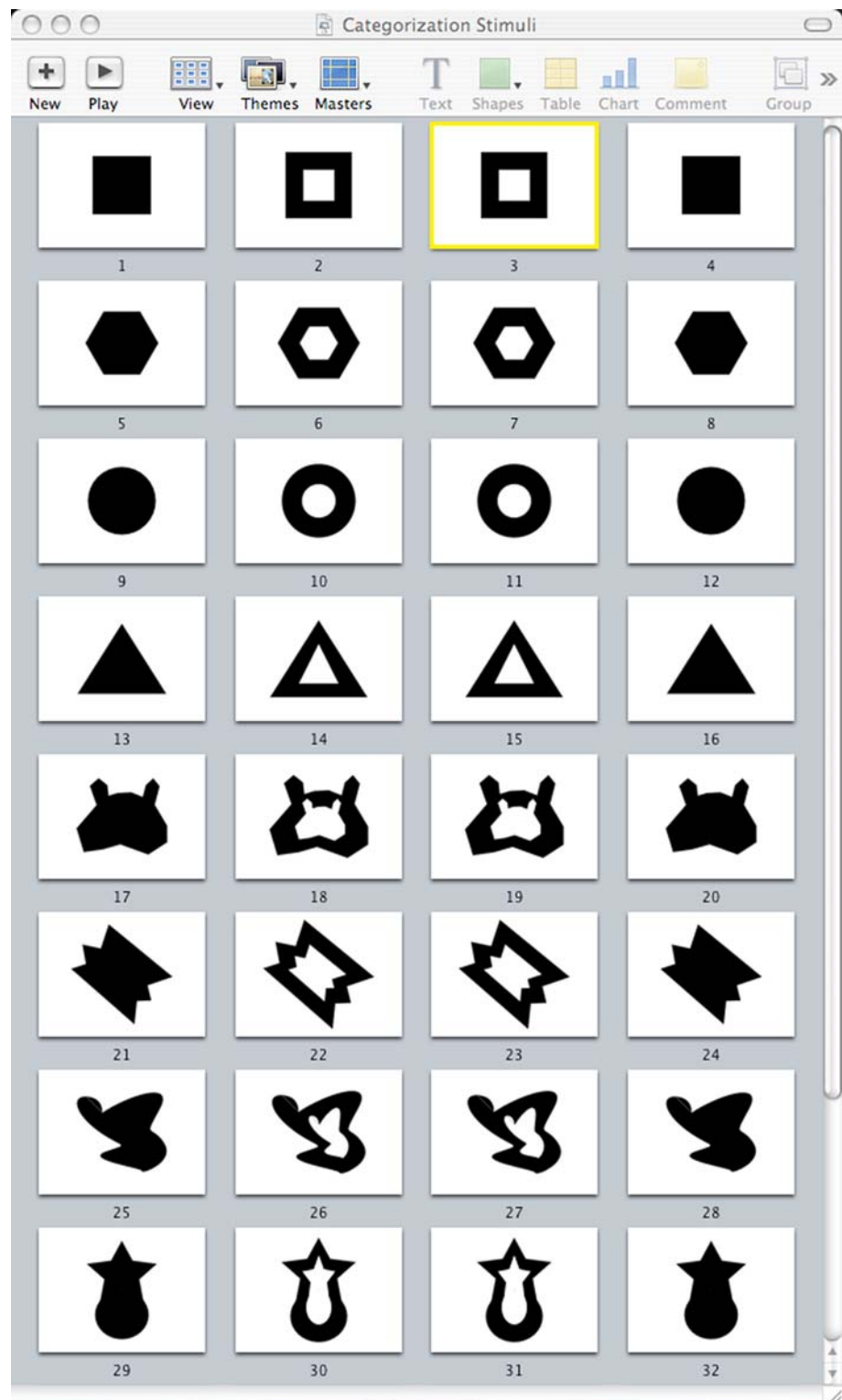
For nonhuman animals, concept learning is considered the highest degree of abstraction attainable (Reichmuth Kastak and Schusterman 2002). Concepts may be centered not only on perceptual properties but also on common relational or associational properties (Roberts 1998). When stimuli are categorized together through common associations with other stimuli or events associative concepts are formed. Relational concepts, on the other hand, result from the learning of abstract relationships that are shared by sets of stimuli. In this respect, relative class concepts (e.g., bigger, darker, etc.) are a measure of advanced cognitive ability because learning to respond relatively is more complex than responding absolutely, as would occur when responding only to the color green (Pepperberg and Brezinsky 1991; Hanggi 2003). This is not to say that understanding relative class concepts is as advanced as other relational concepts. A more complex level would include conditional relations (if A then B) or relations between relations (same versus different). Relative size concepts are demonstrated when an animal learns a specific rule regarding the relationship of size between stimuli. Memory should play a critical role in any concept use because the development of a concept is of little value if it cannot be recalled when needed.

In Experiment 3, we investigated equine LTM for concept use and behavior toward novel stimuli based on a concept of relative size not tested in more than 7 years (Hanggi 2003).

Original procedure

The third LTM test was based on a relative size concept experiment described in detail in Hanggi (2003). The apparatus and testing protocol used in the original procedure

Fig. 3 Open-center and solid stimuli used to test LTM for categorization with Coco Bean in Experiment 2. The first *four rows* depict familiar stimuli (square, hexagon, circle, and triangle) taken from the original experiment while the last *four rows* depict novel stimuli (hippo, concertina, inkblot, and star-drop) presented during the LTM experiment. As in Fig. 1, each stimulus set appears twice per row to facilitate stimulus placement on the LCD multi-displays



were the same as that described in the original procedure for Experiment 1 of this paper. The stimuli were black on white 2D computer-generated images of assorted shapes (circles, squares, triangles, clovers, trees, and several different silhouettes of figures including coyotes, skiers, campfires,

clowns, etc.) as well as colored 3D objects consisting of balls, plastic plates, different shaped PVC connectors, and flowerpots. Four different sizes (large, medium, small, and tiny) of stimuli made up each set and, for Tequila, the larger of any two presented stimuli was always the correct choice.

Table 3 Tequila's performance on long-term memory tests for relative size concept using familiar (circle, clover, tree, campfire) and novel (pot, gargoyle, lobster, beets, basket) stimulus sets

Familiar sets (S+/S−)	CR	Total	% CR	Binomial (z)
Large/medium	14	20	70.0	1.789 ^b
Large/small	20	20	100.0	4.472 ^a
Large/tiny	20	20	100.0	4.472 ^a
Medium/small	17	20	85.0	3.130 ^a
Medium/tiny	20	20	100.0	4.472 ^a
Small/tiny	17	20	85.0	3.130 ^a
Novel sets (S+/S−)	CR	Total	% CR	Binomial (z)
Large/medium	18	20	90.0	3.578 ^a
Large/small	19	20	95.0	4.025 ^a
Large/tiny	20	20	100.0	4.472 ^a
Medium/small	20	20	100.0	4.472 ^a
Medium/tiny	20	20	100.0	4.472 ^a
Small/tiny	19	20	95.0	4.025 ^a

CR correct response, % CR percent correct, S+ positive stimulus, S− negative stimulus

^a $\alpha = 0.01$, ^b $\alpha = 0.05$, binomial test

Another horse was tested for the same relative size concept (larger as correct), while a third was tested for the opposite concept (smaller as correct).

The training phase, comprising a subset of the stimulus sets, was followed by the concept testing and transposition phase. This phase involved more complex novel 2D shapes as well as novel 3D objects. For the transposition tests, a limited number of trials per stimulus set were run in order to minimize the chance that the horses would become too familiar with any given stimulus. After trial-and-error learning with numerous errors during the first discrimination (learning criterion of 20/20 correct consecutive responses, $z = 4.47$, $P < 0.00001$, $\alpha = 0.01$, binomial test), the horses learned new discriminations more rapidly and then consistently chose the correct stimuli, even in novel situations.

The results of this study showed, for the first time, that horses could solve discriminations and transpositions based on relative size concepts. Moreover, they could readily generalize this concept to dimensions outside of their training and testing experience. This demonstrated that horses are capable of making relative judgments about objects and events in their environment and can use basic concepts for problem solving and to facilitate learning. This experiment was last run in 2000.

Procedure for LTM test

Seven years and 4 months after the original study, Tequila's LTM was tested for relative size concept using

familiar sets of stimuli and then novel stimulus sets. The LTM experiment was conducted using the new LCD display apparatus, which provided information about Tequila's ability to generalize between testing systems (Fig. 2b). The four familiar stimulus sets tested were circle, clover, tree, and campfire with 20 trials for each size combination. Sessions comprised 30–40 trials per day and were balanced for the placement of the correct stimulus. This was followed by sessions testing five novel sets of stimuli with 20 trials per size group. These stimuli were taken from the computer application, Photo Objects 150,000 (Art Explosion), color modified so that they displayed as black on white images, and labeled pot, gargoyle, lobster, beets, and basket (Fig. 4).

Results and discussion

Tequila's performance on the familiar sets was well above chance at $\alpha = 0.01$ ($P \leq 0.001$, binomial test) for all size combinations except familiar large versus medium, which was significant at $\alpha = 0.05$, $P = 0.037$ (Table 3). Scoring slightly lower on this size combination was not unusual: research with other animals has shown that some members of a category are classified readily while others are not (Pearce 1994). Tequila's overall performance indicated that he recalled and applied the relative size concept rule that he had learned more than 7 years earlier. This was further supported by his behavior of near perfect stimulus selection when exposed to size trials with novel stimuli, scoring 90–100% correct on the five sets. In addition, he chose correctly on all 15 Trial 1 combinations (large vs. medium, medium vs. small, small vs. tiny) for these new sets. This further confirmed that Tequila applied the concept he remembered from years ago. This experiment, as well as Experiment 2, demonstrates that once a conceptual task is learned, testing may be halted and then resumed months to years later without significant memory decay.

General discussion

Similar to primates and one sea lion, the horses in these experiments demonstrated that they could remember relatively complex problem-solving strategies for a minimum of 7 years and as long as or longer than 10 years and draw on them to work out new challenges of a comparable nature. The use of concepts not only facilitates problem solving but also appears to aid memory as well. When attempting to recall a series of unrelated discriminations, an animal must remember each instance of which stimulus was correct and which was not. When operating under concepts, the animal need only retain a memory of a single

Fig. 4 Familiar and novel stimuli used to test LTM for a relative size concept with Tequila in Experiment 3. Each row depicts four different sizes for which the larger of any two was always the correct choice. The first four rows show familiar stimuli (circle, clover, tree, and campfire) from the original experiment while the last five rows show novel stimuli (pot, gargoyle, lobster, beets, and basket) used during the LTM experiment. Because of the many possible presentation permutations, stimulus sets are ordered here according only to size and shape, not as they appear in the Keynote window during actual testing. During testing, the Keynote window appears similar to those shown in Figs. 1 and 3



rule, which is more economical and less taxing from a cognitive standpoint. The durable nature of the categories and concepts formed years earlier by these horses reveals the cognitive potential of the species with respect to memory. Furthermore, these experiments augment the very limited information available on LTM for categories and concepts in nonhuman animals.

The use of a computer with LCD multi-displays proved highly successful and a first in cognition research with horses. Equine visual acuity is not quite as good as seen in humans with normal vision (Timney and Macuda 2001) and their color vision is similar to humans with red–green color deficiencies (Hanggi et al. 2007). Nevertheless, the horses' immediate adaptation to the LCD displays alleviated our concerns that they might not perceive stimuli on computer screens well enough for research purposes.

In order to live in a functional social system as horses do, individuals must be capable of recognizing each other and remembering relationships. Social dominance occurs whenever two or more horses are together and individuals establish relationships within an interactive system based on rank order, social attachments, and antagonistic relations (Kolter and Zimmermann 1988). Social cognition is beneficial for life in social groups, e.g., one horse drawing conclusions about its own social status from observing the interactions of conspecifics can facilitate its integration into that group (Krueger and Heinze 2008). Associations and rules within a herd and between herds must be learned and then remembered often over periods of years. Horses generally form social groups of 20 individuals or less and it is suggested that this number represents the mental limit of a horse's social memory (Mills and Nankervis 1999). Larger herds are generally comprised of smaller bands that coexist as transitory social groups. Upon close examination, each band segregates itself within the larger group and individuals from different bands display more caution or aggression toward one another. Social units that are temporarily separated are able to regroup as well as return to previously known home ranges (Keiper 1979).

That horses develop and learn when young and remember important experiences later on in life has significant implications for the human/horse relationship and equine management. What is learned early in life frequently remains part of the horse's memory for an indefinite period and, if negative, can have serious consequences for its well-being. For example, "imprint training" (a misnomer) of foals gained popularity over the past 15 years with many owners attempting to desensitize newly born foals to a variety of objects and events (general handling, farrier and veterinary work, unusual objects) that they would encounter over the course of their lives (Miller 1991). In theory, exposing foals to an assortment of stimuli during an early learning period could provide benefits. Due to their size,

foals are easier to handle than larger horses, and desensitizing them early on should make subsequent training easier as well. However, a number of studies on the short- and long-term benefits of early, intensive handling of foals found no profound differences in response to handling between foals subjected to regimented handling immediately after birth and those handled under routine practices (see Diehl 2005, for a review). Moreover, foals may not perceive forced human tactile stimulation as positive (Henry et al. 2006). Although desensitization is not on the same level as higher-order cognitive abilities, memories of such horse–human encounters may be recalled for a very long time. Extreme caution, therefore, as well as a thorough understanding of desensitization training, is required when handlers attempt any sort of training of very young horses. Horses handled improperly at an early age may later exhibit fear responses to stimuli and events that otherwise would have been perceived benign (Hanggi, unpublished data). More research is certainly needed into the long-term effects of this type of training.

It is evident that horses depend on LTM to manage an assortment of cognitive, social, and ecological problem-solving situations. The 6- to 10-year retention periods documented here demonstrate that memories endure for a considerable time within a horse's life span. A better understanding of equine cognitive capabilities with respect to categorical and conceptual behavior and LTM should lead to improvement in training and management. This should benefit not only horses that interact with humans but those that live wild as well.

Acknowledgments This research was funded by grants and public donations to the Equine Research Foundation (ERF). Thank you as always to Meinrad and Betty Hanggi for their encouragement and support. Donations and grants from Adobe Systems/Gifts In Kind International, ASPCA, Claire Giannini Fund, Compumentor, Fleet Matching Gift Program, Microsoft and Troxel greatly helped the ERF research, as did the wholehearted involvement of interns and ERF learning and riding vacation participants. We also thank the editor and referees for their valuable suggestions. The article, "Optimizing animal cognition testing protocol using the Equine Research Foundation Training System™ (ERTS)" is available directly through the Equine Research Foundation. Equine care, training, and testing procedures used in these experiments comply with the current laws of the United States of America.

References

- Burdyn LE Jr, Noble L, Shreves LE, Thomas RK (1984) Long-term memory for concepts by squirrel monkeys. *Physiol Psychol* 12:97–102
- Crystal JD, Shettleworth SJ (1994) Spatial list learning in black-capped chickadees. *Anim Learn Behav* 22:77–83
- Diehl NK (2005) Review of research on the effectiveness of early intensive handling of foals. In: AAEP 51st Annual Convention Proceedings. Seattle, Washington, pp 246–255

- Dixon J (1970) The horse: a dumb animal? ... Neigh! *Thoroughbred Rec* 192:1654–1657
- Fagot J, Cook RG (2006) Evidence for large long-term memory capacities in baboons and pigeons and its implications for learning and the evolution of cognition. *PNAS* 103(46):17463–17567
- Fiske JC, Potter GD (1979) Discrimination reversal learning in yearling horses. *J Anim Sci* 49:583–588
- Giebel HD (1958) Visuelles Lernvermögen bei Einhufern. *Zool Jb* 67:487–520
- Grzimek B (1949) Gedächtnisversuche mit Pferden. *Z Tierpsychol* 6:445–4454
- Hanggi EB (1999a) Categorization learning in horses (*Equus caballus*). *J Comp Psychol* 113(3):243–252
- Hanggi EB (1999b) Interocular transfer of learning in horses (*Equus caballus*). *J Equine Vet Sci* 19(2):518–523
- Hanggi EB (2001) Can horses recognize pictures? In: Proceedings of the third international conference of cognitive science, Beijing, China, pp 52–56
- Hanggi EB (2003) Discrimination learning based on relative size concepts in horses (*Equus caballus*). *Appl Anim Behav Sci* 83:201–213
- Hanggi EB (2005) The thinking horse: cognition and perception reviewed. In: American Association of Equine Practitioners 51st Annual Convention Proceedings. Seattle, Washington, pp 246–255
- Hanggi EB (2006) Equine cognition and perception: understanding the horse. In: Fujita K, Itakura S (eds) Diversity of cognition: evolution, development, domestication and pathology. Kyoto University Press, Kyoto, pp 98–118
- Hanggi EB (2007) Optimizing animal cognition testing protocol using the Equine Research Foundation Training System™ (ERTS). Equine Research Foundation Publications, CA, pp 1–5
- Hanggi EB, Ingersoll JF, Waggoner TL (2007) Color vision in horses (*Equus caballus*): deficiencies identified using a pseudoisochromatic plate test. *J Comp Psychol* 121(1):65–72
- Harlow HF (1949) The formation of learning sets. *Psychol Rev* 56:51–65
- Henry S, Richard-Yris M-A, Hausberger M (2006) Influence of various early human-foal interferences on subsequent human-foal relationship. *Dev Psychobiol* 48(8):712–718
- Huber L (1995) On the biology of perceptual categorization. *Evol Cogn* 1(2):121–138
- Johnson CK, Davis RT (1973) Seven-year retention of oddity learning sets in monkeys. *Percept Mot Skills* 37:920–922
- Keiper RR (1979) Population dynamics of feral ponies. In: Waring GH (ed) *Horse Behavior*, 2nd Edition edn. Noyes, New York, p 237
- Kolter L, Zimmermann W (1988) Social behaviour of Przewalski horses (*Equus p. przewalskii*) in the Cologne Zoo and its consequences for management and housing. *Appl Anim Behav Sci* 21:117–145
- Krueger K, Heinze J (2008) Horse sense: social status of horses (*Equus caballus*) affects their likelihood of copying other horses' behavior. *Anim Cogn* 11:431–439
- Linklater WL (2007) Equine learning in a wider context—opportunities for integrative pluralism. *Behav Proc* 76:53–56
- Marinier SL, Alexander AJ (1994) The use of a maze in testing learning and memory in horses. *Appl Anim Behav Sci* 39:177–182
- McLean AN (2004) Short-term spatial memory in the domestic horse. *Appl Anim Behav Sci* 85:93–105
- Miller RM (1991) Imprint training of the newborn foal. Western Horseman, Inc, Colorado
- Mills D, Nankervis K (1999) Equine behaviour: principles & practice. Blackwell Science, Oxford, p 129
- Murphy J, Arkins S (2007) Equine learning behaviour. *Behav Proc* 76:1–13
- Nobbe DE (1978) Delayed response in the horse. In: Waring GH (ed) *Horse behavior*. Noyes, New York, p 113
- Patterson RL, Tzeng OJL (1979) Long-term memory for abstract concepts in the lowland gorilla (*Gorilla g. gorilla*). *Bull Psychon Soc* 113:279–282
- Pearce JM (1994) Discrimination and categorization. In: Mackintosh NJ (ed) *Animal learning and cognition*. Academic Press, San Diego, pp 109–134
- Pepperberg IM, Brezinsky MV (1991) Acquisition of a relative class concept by an African Gray parrot (*Psittacus erithacus*): Discriminations based on relative size. *J Comp Psychol* 105(3):286–294
- Reichmuth Kastak C, Schusterman RJ (2002) Long-term memory for concepts in a California sea lion (*Zalophus californianus*). *Anim Cogn* 5:225–232
- Roberts WA (1998) Principles of animal cognition. McGraw-Hill, Boston
- Sappington BF, Goldman L (1994) Discrimination learning and concept formation in the Arabian horse. *J Anim Sci* 72:3080–3087
- Timney B, Macuda T (2001) Vision and hearing in horses. *J Amer Vet Med Assoc* 218(10):1567–1574
- Vauclair J (1996) *Animal cognition*. Harvard, Cambridge, pp 18–20
- Waring G (2003) *Horse behavior*, 2nd edn. Noyes, New York
- Warren JM, Warren HB (1962) Reversal learning by horse and raccoon. *J Genet Psychol* 100:215–220
- Zayan R, Vauclair J (1998) Categories as paradigms for comparative cognition. *Behav Proc* 42:87–99