VISUAL DISCRIMINATION AND OBJECT/PICTURE RECOGNITION IN HENS.

A thesis submitted in fulfilment

of the requirements for the degree of

Doctor of Philosophy

at

The University of Waikato

by

RENÉE CARON RICHARDS RAILTON

2011

ABSTRACT

Eight experiments were conducted to examine different aspects of hen's visual behaviour, and to assess whether hens responded to photographs in the same way they do to the real objects that were depicted in the photographs. In Experiment 1, six hens were trained to perform either a conditional discrimination (successive) or forced-choice discrimination (simultaneous) between flickering (25 Hz) and steady lights. A descending method of limits procedure was then used to increase the flicker speed by 5 Hz over blocks of 20 trials until percentages correct decreased below 55%. The critical flicker fusion frequency of hens was found to range between 68.5 and 95.4 Hz (at a luminance of 300 cd/m^2). In Experiment 2, hens were trained to discriminate between steady images presented on a TFT screen, and tested for transfer of that discrimination to a CRT monitor at different refresh rates, on which the images were assumed to appear flickering. It was found that hens showed transfer across all refresh rates with coloured stimuli, but that the degree of transfer decreased as refresh rate decreased with stimuli that were discriminable only on shape. In Experiment 3, a similar decrease in accuracy was shown as refresh rate decreased using a range of stimuli. However, hens did not learn to discriminate all stimuli, and thus transfer could not be assessed with some stimuli. Experiment 4, hens were trained with flickering images and showed relatively high transfer to less flickering, or steady, images. In Experiment 5, a procedure was developed to assess whether hens transferred a discrimination of 3D object to 2D photographs of those objects, and vice versa. In Experiment 6, hens were trained to discriminate stimuli of different colours, or of different shapes. The hens learned to discriminate, and transferred this discrimination, with the coloured shapes. The hens also learned to discriminate the same colour (but differently shaped) stimuli, however, further testing showed that an extraneous variables had come to control behaviour. As a result, the equipment was modified for Experiments 7 and 8. In both experiments, only three of the six hens showed discrimination to any degree, and none transferred this discrimination to photographs or objects. It was concluded that hens do not respond to objects depicted in pictures in the same way they do to the real objects. Thus, these experiments show that that animals' visual systems need to taken into account when visual stimuli are used in research, and researchers first need to establish that animals can see the visual stimuli and that the method of stimulus

presentation is species appropriate if images are to be used as representatives of real world stimuli.

ACKNOWLEDGEMENTS

My deepest thanks and appreciation goes to Professor Mary Foster and Associate Professor Bill Temple for their guidance, encouragement and support throughout this thesis. Thank you both for steering me through this process with the utmost patience.

Special thanks to Jenny Chandler and Rob Bakker for your technical support, expertise and friendship. Thank you also to the members of the Animal Behaviour and Welfare Research Centre for running my experiments for all these years, and for your support and advice.

Finally, extra special thanks to my friends – and especially to Amy, Aimee and Jennifer, for all your proof reading brilliance. You all have no idea how much you have helped me.

DEDICATION

This thesis is dedicated to my mother, Gerardine Hura. Thank you for always believing in me, and for knowing why I needed to do this. You could see the reasons why, when sometimes I could not.

CONTENTS

PAGE

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

Humans rely on the use of two-dimensional images to represent three-dimensional objects. Such images can vary from highly stylised cartoon-like representations to detailed photographs and moving images. Two-dimensional pictorial images have also been used widely in animal research based on an assumption that these pictures accurately represent the object to the animal. For example, researchers have used two-dimensional images to examine animals' visual recognition (e.g., Candland, 1969; Weavers, 2000), concept formation and categorization (e.g., Herrnstein & Loveland, 1964; Herrnstein, Loveland & Cable, 1976, Cerella, 1979; Herrnstein & de Villiers, 1980; Greene 1984), spatial memory learning (e.g., Spetch & Wilkie, 1994), social facilitation (e.g., Keeling & Hurnik, 1993), and motion perception (e.g., Lea & Dittrich, 1999), and video images have been used to measure social behaviour (i.e., responses to conspecifics, e.g., Clark & Uetz, 1990; Evans & Marler, 1991).

Pictorial stimuli can range from simple abstract stimuli to more complex representational stimuli. Abstract, or non-representational stimuli (such as geometric forms), have often been used to assess animal perception. For example, DeMello, Foster and Temple (1992) used grey gratings presented behind stimulus keys to examine visual acuity in hens. Zonderland, Cornelissen, Wolthuis-Fillerup and Spoolder (2008) used Landolt C symbols to assess visual acuity in pigs. Brodbeck and Shettleworth (1995) used different colours presented on a colour monitor to assess spatial cues used by chickadees and junco birds.

More complex stimuli such as photographs or videos are often used as if they have properties related to those of the physical object and can be used as a substitution of that object. Parron, Call and Fagot (2008) assessed baboons' ability to recognise photographs as representing objects. The authors used photographs of a food item (banana) and photographs of a non-food item (pebble) to assess behavioural responses to the photographs. They found that the baboons selected the picture of the banana more often than that of the pebble, but that many of the baboons ate the photographs of the banana showing that they confused the photographs for the real thing. Interestingly, the authors found that chimpanzees and gorillas did not show the same level of object/picture confusion. Other studies have used pictorial stimuli as if they represent the real stimuli. For example, Candland (1969) used coloured slides of roosters to assess conspecific recognition in chickens. He concluded that the rooster's combs were most easily discriminated by the chickens. This use of images assumes that the slides are functioning as a substitution of the real animals; however, this was never assessed by Candland. While there is extensive use of pictorial stimuli in behavioural research, there is relatively little research on picture-object correspondence (Spetch, Kelly & Reid, 1999).

A range of media have been used to present visual stimuli to animals. These can range from still stimuli such as line drawings, slides, and photos, to moving stimuli such as computer and television playback. D'Eath (1998) and Bovet and Vauclair (2000) reviewed the use of two-dimensional images in animal research. They pointed out that the use of still and moving pictorial stimuli have a number of advantages in research. Such images give the experimenter greater control over the presentation and arrangement of the stimuli than would be possible with real animals or objects. They allow repeated presentation of identical stimuli to several subjects, reducing the variability that can arise through direct interaction with the stimuli. In addition, the use of images reduces the amount of handling and number of subjects required (Ophir & Galef, 2003). Furthermore, they allow researchers to manipulate the images in a manner that allows assessment of the aspect of the picture to which the animal is attending. These manipulations may be impossible with the real stimulus or object. For example, Candland (1969), using slide stimuli, assessed what features chickens used in conspecific recognition by swapping the combs and beaks of different roosters; something that is biologically impossible in real life.

Although there are a number of advantages with the use of pictorial stimuli, there are also disadvantages. As previously pointed out, using artificial stimuli relies on the assumption that the animal can see these images, and will respond to them in the same way they would if the real stimuli were present. Other aspects of animals' visual systems need to be taken into account. In the case of television and computer monitors it is necessary that the animal sees the images as fused into motion rather than as a series of static images. In addition, recognition of visual stimuli can be affected by the lack of depth cues, the animals' visual acuity, and the absence of interaction with the object they are viewing (Watanabe & Troje, 2006).

In humans, it is generally learned from a young age that pictorial stimuli are representations of real stimuli, but are not the real stimuli, and recognition of objects in pictures can be difficult for cultures that have had little exposure to pictures. Miller (1973) outlined some studies showing cross cultural differences in ability to recognise images in photographs. For example, Kidd (1904; cited in Miller, 1973) found that Bantus were unable to perceive objects in photographs until the details of those objects were pointed out. In another example, Deręgowski, Muldrow and Muldrow (1972; cited in Deręgowski, 2000) found that the Mekan (a remote population in Ethiopia) could not recognise pictures drawn on paper. However they could learn to recognise the pictures when they were drawn on coarse cloth (which the Mekan had experience with) and after the experimenters had pointed out the features of the pictures. These studies suggest that humans do not show correspondence between pictures and real stimuli without some degree of experience and/or training and implies that animals may also require training.

It is reasonable to assume that abstract or non-representational stimuli (e.g., geometric forms, gratings, Landolt C) will function in the intended fashion for animals. However, when the stimuli are intended to represent real objects, that they will function as such is often assumed by researchers without any real evidence. For example, Herrnstein and Loveland (1964) were the first researchers to test for categorisation in birds. They found, using a go-no/go procedure, that pigeons were able to discriminate slides that contained human figures from those that did not. The pigeons generalised this discrimination to novel pictures with or without people. The authors argued that the pigeons were able to categorise the pictures based on whether humans were present or not. Herrnstein et al., (1976) extended these findings and found that pigeons were able to discriminate between images showing the presence or absence of water, trees, or a specific person. Other categorisation studies have shown that pigeons can discriminate between pictures of different types of leaves (Cerella, 1979), underwater pictures with and without fish (Herrnstein & de Villiers, 1980) and between "non-natural" scenes, for example, Monet and Picasso paintings (Watanabe, Sakamoto & Wakita, 1995). However, it is unclear from these studies which feature of the image the pigeons were using in their discrimination. As these studies did not assess if the pigeons responded in a similar manner to the real objects as they did to the photographs, it is unclear what feature they were responding to (Weisman & Spetch, 2010). In another study assessing categorisation, Greene (1984) found that pigeons that had learned to discriminate between pictures with and without the presence of humans were doing so based on features on the background rather than whether humans were present in the pictures. Vaughan and Green (1984) found that pigeons were able to remember a large number of slides (up to 320 pictures) and that performance was not disrupted even after 731 days before retesting, showing that pigeons have the capacity to remember details about a large number of images. The authors argue that any transfer of performance found in concept formation studies may simply be "memorization of particular slides or particular features, coupled with generalization along certain physical dimensions" (p. 270), rather than evidence of categorisation in birds.

The assessment of correspondence between pictures and the real object is generally measured in two ways. One method is to present images of stimuli which would normally produce specific spontaneous responses (e.g., food, conspecifics or prey species) and assess whether animals respond in a similar manner as they would to the real object. For example, Clark and Uetz (1990), using images presented on a television screen, showed that jumping spiders will attack images of prey, will show courtship behaviour to images of conspecifics, and will retreat from images of predators. In addition, pigeons (Shimizu, 1998); tiger barb fish (Clark & Stephenson, 1999); zebra finches (Galoch & Bischof, 2007); jacky dragon lizards (Ord, Peters, Evans, & Taylor, 2002); and stickleback (Rowland, Bolyard, Jenkins & Fowler, 1995) have all been shown to respond spontaneously to various similarly presented images. Such studies suggest that some important aspect of the real object is perceived in the video image and that this then elicits an appropriate response. However, while it may be that there is some similarity between the real stimulus and its video image, researchers must be careful not to assume that they are identical for that animal.

Another method for measuring picture-object correspondence is to train a particular response to real stimuli and test for transfer of that response to pictures of those stimuli, and vice versa. For example, Patterson-Kane, Nicol, Foster and Temple (1997) trained hens to discriminate between different coloured cardboard, a white hen versus no hen, and a brown hen versus no hen. They were then tested with videoed images of the stimuli. The hens showed no transfer of performance to the videoed images when colour cues were not available (i.e., white hens versus no hen). In addition, the hens learned to discriminate between a real brown hen and a real basketball quickly, but required several hundred trials to learn the same discrimination with videoed images. Neither group were successful at transferring their discrimination to the alternative stimuli (i.e., from video to real images, and vice versa). The authors concluded that the videoed images were not equivalent to the real images for the hens.

Correspondence between pictorial stimuli and the real object has been examined using a range of species and methodologies. To name just a few, pigeons transferred their discrimination of spherical/non-spherical objects to photographs and drawings (Delius, 1992), but did not show correspondence between real-life humans and their photographs (Dittrich, Adam, Ünver and Güntürkün, 2010), capuchin monkeys were able to match real objects and photographs/silhouettes/line drawings in a matching-to-sample discrimination (Truppa, Spinozzi, Stegagno & Fagot, 2009), male jacky dragons showed aggressive displays to both live and videoed male conspecifics (Ord et al., 2002), and tiger barb fish showed no difference in schooling behaviour with live, videoed or computer-animated fish (Clark & Stephenson, 1999).

One issue in this research area is that methods used for presenting visual stimuli have been designed with the human visual system in mind, and important cues may be missing for animals. Researchers have to be careful not to ignore the fact that animals can have very different visual systems. For example, humans have three types of cones and can see wavelengths ranging between 380-780 nm. However, birds have at least four cone types and can see wavelengths across a much wider range, including the ultraviolet colour spectrum, which humans cannot (Lewis & Morris, 1998). As visual stimuli do not contain ultraviolet wavelengths, this may distort the images for birds. It is important to first establish whether an animal with its particular visual systems can be expected to respond to a pictorial stimulus as if it were real.

A number of questions arise from research using visual stimuli. Firstly, can animals see and respond to stimuli presented on television and computer monitors as if they are real? As some monitors have a refresh rate, hens' ability to see flicker is assessed in Experiment 1. In Experiments 2 to 4, hens were trained to discriminate between various pictorial stimuli on either a cathode ray tube (CRT) or thin film transistor (TFT) monitor. The flicker rate of the CRT monitor was then altered to assess how flicker affected the hens' performance. Secondly, do simple images (i.e., photographs) actually function as a substitution of the real objects for hens? In Experiments 5 to 8, hens' correspondence between photographs and the real object was assessed.

EXPERIMENT 1[1]

Visual systems are adapted to seeing moving stimuli (Lea & Dittrich, 1999) and all visual systems have a rate at which a flashing or intermittent stimulus is fused and seen as steady. The lowest frequency at which a flickering light source is seen as steady or continuous is defined as critical flicker fusion (CFF) (Brundett, 1974; Landis & Hamwi, 1954). Presenting a series of static images at a rate higher than the CFF threshold of an organism will lead the images to appear as continuously moving (D'Eath, 1998). The CFF threshold is the frequency at which flicker is detected 50 % of the time. CFF can be affected by stimulus luminance, colour, size, its retinal position, and by fatigue in the observer (Ginsberg & Nilsson, 1971; Nuboer, Coemans & Vos, 1992; Ricciuti & Misiak, 1954).

Flicker fusion is necessary for technology that presents moving images. Cathode ray tube (CRT) monitors present static images in rapid succession so that the images appear to be moving. This is done by an electron gun which lights individual phosphor dots at a certain rate, termed a refresh rate. In between the electron scans, these phosphor dots darken. If the rate at which these images are presented falls below the CFF of an organism, the flicker will become apparent. As a result, refresh rates are commonly set at 50-85 Hertz (Hz) which is above the CFF of human visual systems (reported to be between 50-60 Hz (Brundett, 1974; Hart, 1992)). Accordingly, the human visual systems will fuse images presented at a frequency above their CFF and see the images as continually present and moving.

Television and computer monitors have often been used to present stimuli to animals (e.g., Bradshaw & Dawkins, 1993; Clark & Stephenson, 1999; Troje, Huber, Loidlt, Aust & Fieder, 1999). However, CFF values vary across species. For example, spiders have CFF values in the range of 10-37 Hz (DeVoe, 1963; Muñoz-Cuevas, 1984), and geckos and horned lizards have CFF values ranging between 25 and 42 Hz (Crozier & Wolf, 1939 & 1941, cited in Fleishman, Marshall & Hertz, 1995). Fast flying insects on the other hand have CFF values above 200 Hz (Autrum, 1950; Laughlin, 1981). Pigeons' CFFs have been reported to range between 65 and 145 Hz (Hendricks, 1966; Powell, 1967), and hens' to range between 69 and 105 Hz

---

[1] This experiment has been published as Railton, R.C.R., Foster, T.M., & Temple, W. (2009). A comparison of two methods for assessing critical flicker fusion frequency in hens. *Behavioural Processes, 80,* 196-200.

(Nuboer et al., 1992). Researchers must be careful to present images using a medium that is higher than an animal's CFF. Any animal with a CFF higher than the particular monitor that is used to present the images may see them as flickering and distorted.

Some species do respond to stimuli presented on television or computer screens as if they were real. Clark and Uetz (1990) presented images to jumping spiders on a CRT screen set at 60 Hz. They found that the spiders responded appropriately (approached, attacked, or showed courtship behaviours) to images of prey, predators and conspecifics. They did not discriminate between live prey and their simultaneously presented video images showing that the flicker rate of the screen did not impact on the spiders' discrimination. Pigeons (Shimizu, 1998); tiger barb fish (Clark & Stephenson, 1999); zebra finches (Galoch & Bischof, 2006); jacky dragon lizards (Ord et al., 2002); and stickleback (Rowland et al., 1995) have all been shown to respond appropriately to various similarly presented images.

However, a number of studies with birds have failed to train successful discrimination between stimuli when they have been presented as video or computer images. Patterson-Kane et al. (1997) found that hens had difficulty learning to discriminate between two sets of video images of other hens, and that once the discrimination was acquired it did not generalize to the real objects. Similarly, although the hens learned to discriminate between the real objects easily, they were unable to generalize this discrimination to video images of the objects (except when the discrimination could be based on colour alone). These data suggest that video images were not equivalent to the real object for these hens. It may be that the refresh rate of the computer monitor used was below the CFF for the hens. D'Eath and Dawkins (1996) found hens began feeding more quickly near a familiar rather than unfamiliar hen; but showed no discrimination when this was repeated with life-size colour CRT video images. Ryan and Lea (1994) found that pigeons did not respond to life-size moving video images of conspecifics.

In the few studies that report apparent success, other factors may have contributed to the successful discrimination. For example, McQuoid and Galef (1993) found that Burmese fowl acquired food dish preferences after observing videos of conspecifics eating out of marked dishes; however, the authors suggest that these preferences could have been a result of vertical movement in the vicinity of a food dish rather than the specific sight of a conspecific feeding at that dish.

Troje et al. (1999) found pigeons were able to discriminate between images of male and female faces presented on a computer monitor, however, they concluded that pigeons were discriminating based on the luminance rather than the shape of images. Adret (1997) measured male zebra finches approach and singing towards videos of various stimuli. She found that conspecific finches were approached more often and induced more singing than other stimuli; however, images of fish, locust, and computer animated doodles had a higher approach score than a conspecific with white plumage or a lovebird. The author notes that the finches were most likely responding to the amount of movement in the images, rather than recognition of the images. Evans and Marler (1991) found that cockerels will increase their production of alarm calls in the presence of a video image of a hen; similar to levels shown with live hens. However, the images were accompanied by a sound track of the hen which may have been the reason for this positive result. It may be that the sound track on its own was enough to produce these results.

A few studies have successfully used video playback with birds. Bird and Emery (2008) found that rooks preferred video images of a familiar conspecific over a non-familiar conspecific indicating that rooks can recognise individuals in the videoed images. Toda and Watanabe (2008) trained pigeons to peck at live video images of themselves, rather than pre-recorded videos, indicating they could discriminate between the two video types. These two studies used TFT/LCD screens to present the stimuli. These types of screens are virtually flickerless and may be easier for animals with a high CFF threshold to view. This will be discussed in more detail later in this thesis.

One possible reason for some of the failures of birds to discriminate video stimuli could be that their CFF may be higher than the refresh rate of the monitors. Two studies using different procedures have examined CFF in hens, and both suggest that hens have higher CFF values than humans (Jarvis, Taylor, Prescott, Meeks & Wathes, 2002; Nuboer et al., 1992). Jarvis et al. compared the fit of two models to measures of flicker sensitivity for hens and humans. The flicker sensitivity of the hens was assessed by requiring them to select the flickering light from a pair (one flickering, one steady). This simultaneous choice procedure is termed two-alternative forced choice (2AFC). The hens generally had higher flicker sensitivity values than humans across a range of luminance levels.

Nuboer et al. (1992) examined CFF in two hens using a conditional discrimination procedure requiring a successive discrimination between a flickering and a steady light. Hens were trained to move toward one or the other side of a chamber and break a beam to indicate whether the light was steady or flickering in a trial-by-trial staircase procedure, where the flicker rate was increased or decreased after each correct or incorrect response respectively. They found the maximum CFF was 105 Hz using a monochromatic light (476 nm), but was less for the other wavelengths used.

While both studies show that hens have far higher CFF thresholds than humans, Jarvis et al. (2002), using a simultaneous presentation task, found lower CFF values overall than Nuboer et al. (1992) had, using a successive discrimination task. This may be a result of the different procedures used in the two studies. It would therefore be useful to determine if there is a difference in CFF values found with these two methods of stimulus presentation.

It has been suggested that successive and simultaneous presentation give different effects. McLean and White (1983) found discrimination was more accurate and stable when red and green stimuli were presented simultaneously rather than successively. Ulrich and Miller (2004) note that 2AFC (simultaneous) procedures are often preferred in studies determining thresholds as they lead to lower response biases and better performances than the classic successive procedure. However, Bushnell (1999) found that rats' detection of visual stimuli was more accurate and stable with successive than with simultaneous procedures. Furthermore, Shelton, Picardi and Green (1982) found little to no effect on auditory threshold performance across these methods of stimulus presentation. Siegel and Honig (1970) trained pigeons to discriminate between slides either with or without the presence of a human. They used both successive and simultaneous presentation of images and found that the pigeons learned the discrimination at a similar rate across the two training procedures and were able to transfer the task across presentation methods.

The present study aimed to investigate CFF with hens using a 2AFC procedure and a conditional discrimination procedure, at one particular luminance level. The aim was to (a) produce more data on CFF with hens; and (b) establish whether the differences in CFF values found between Jarvis et al. (2002) and Nuboer et al. (1992) were a result of the different procedures they used.

Method

Subjects

Six experimentally naïve flock-reared Brown Shaver-Starcross hens (numbered 21-26) served as subjects. The hens were 18 months old at the beginning of the experiment. They were individually housed in metal cages (300-mm high × 450-mm long × 450-mm wide) in a ventilated room that was lit on a 12:12-h light:dark cycle with two 100-W incandescent lights. Water was freely available, and grit and vitamins were supplied weekly. Throughout the experiment all hens had red fleshy combs suggesting good health. The hens were weighed daily and provided with supplementary feed (commercial laying pellets) if required to maintain them at approximately 80 % (+/–5 %) of their free-feeding body weights. The principles of laboratory animal care were followed and all procedures were approved by the University of Waikato Animals Ethics Committee for all experiments in this thesis.

Apparatus

The experimental chamber (410-mm wide × 580-mm long × 540-mm high) was made of 20-mm thick white particle-board. The chamber floor was covered with a thick metal grid (30 mm × 30 mm) enclosed in a removable steel tray. Two white back-lit response keys (30-mm diameter) were positioned on the right hand wall of the chamber 400-mm high and 200-mm apart. Behind each response key, the stimulus was a white single chip LED with a typical brightness of 300 micro-candelas. Each response key was surrounded by a metal plate (70-mm wide × 140-mm high). A brief audible feedback beep sounded when a response was made on either lit key. Centrally located below these response keys was an aperture (70-mm wide × 100-mm high) that allowed access to a magazine containing wheat. The hens were observed to view the stimulus keys from a distance of between 5 and 20 cm. A computer (Dell Optiplex GXa) controlled the experimental equipment and recorded all data using Med-PC® software (Version 4). Total session data were also manually recorded into a data book at the end of each session.

The different rates of flicker of the stimulus were produced by a microcontroller (AT90S2313) clocking at 10Mz and interrupted every 2 µs to update a timer. When the timer reached its terminal count an output was toggled, the timer's initial value was reloaded, and the process was repeated. The output was connected to a constant current LED driver. The initial timer value was passed to the micro controller from the Med PC control program. The frequencies of the flicker rate for the diode were calibrated against the frequency counter function of 'FLUKE 79 Series MultiMeter' which has a resolution of 0.1 Hz. A dark:light ratio of 1:1 was used.

Procedure

Prior to the start of discrimination training, the hens were magazine trained and their behaviour was shaped, through successive approximations, to peck a lit key in a single-key chamber. Once the hens were pecking the key reliably, discrimination training began.

Discrimination training.

Successive presentation. Each session began when the two key lights were lit. Both key lights were either flickering (25 Hz) or steady for each trial. When the key lights were flickering the correct response was to peck the left key, and when both key lights were steady, the correct response was to peck the right key. Whether the two stimulus lights were flickering or steady was controlled pseudo-randomly, according to a predetermined series, to ensure that no more than three flickering or steady trials occurred consecutively and that the number of each type of trial would be approximately equal within a session. A variable ratio (VR) schedule of reinforcement was in effect. The VR requirement was gradually increased over 22 sessions until all hens were responding on a VR 10 schedule of reinforcement. This meant that only one in ten correct responses was followed by a reinforcer.

After every correct response, the response keys were turned off and a 1-W white bulb illuminated the magazine for a 3-s period, and, if a reinforcer was scheduled to occur, the hen was provided access to the magazine for 3 s. When an incorrect response was made the response keys were turned off for 3 s. There was a 3-s intertrial interval before the next trial began.

In order to control the number of reinforcers allocated to each of the two stimuli, the computer selected, pseudo-randomly, whether the next trial on which a correct response could result in reinforcement would be a flickering or steady trial. It was not possible for the animal to receive a reinforcer for any other response until the scheduled reinforcer had been collected. Therefore, if the next reinforcer was scheduled for a correct response on a steady trial, any correct responses on flickering trials were followed by the magazine light being turned on, but no food presentation, until the scheduled reinforcer for a correct response on the steady trial had been collected.

Sessions were terminated after 40 min had elapsed or after 30 reinforcers had been obtained, whichever occurred first. Once five sessions at, or above, 90 % correct had been reached threshold trials began. All hens reached this criterion within 28 sessions.

Threshold sessions. A descending method of limits procedure was used during threshold sessions. These began with 20 trials at the training flicker speed (25 Hz) using the normal training procedure. If the percentage of correct responses at the completion of this first block of trials was 90 % or above, the flicker rate was increased by 5 Hz and another block of 20 trials was conducted. If this block of trials gave 55 % correct or greater, another block of 20 trials was started with a further increase of 5 Hz. These increments in flicker rate continued until the percentage of correct responses fell below 55 % for a block of 20 trials. Once this happened, a final 20 trials at the training flicker rate (25 Hz) was conducted before the session was terminated. Any sessions where a hen did not reach 90 % correct in the first or final block of trials were termed unsuccessful threshold sessions and were followed by a normal training session, otherwise another threshold session occurred. Once the hen regained the criterion of 90 % correct or above in a non-threshold session, the next session was a threshold session.

Correct responses were reinforced on an intermittent basis (VR 10) at all flicker frequencies. The magazine light was illuminated after a correct response and a 3-s blackout period followed an incorrect response. The reinforcer rate per trial remained constant across threshold sessions. Threshold sessions had no time limit for completion and each hen received a total of 10 successful threshold sessions.

Discrimination training

Simultaneous presentation. Once the successive threshold sessions were completed, training began with simultaneous presentation, where one key light was flickering and one key light was steady on each trial. Again, the training flickering light was set at 25 Hz. The position of the flickering light was controlled using a predetermined series which ensured that no more than three consecutive trials occurred with the flickering light on the same key. After a correct response (peck the flickering light) the key lights were turned off, the magazine was illuminated for 3 s, and the hen was provided with 3-s access to the magazine if a reinforcer was scheduled. After an incorrect response, the key lights were turned off for 3 s. There was a 3-s intertrial interval before the next trial began. Again the trial type and reinforcement were controlled as in the successive training procedure. The VR requirement was gradually increased over five sessions until all hens were responding on a VR 10 schedule of reinforcement. Sessions were terminated after 40 min had elapsed or after 30 reinforcers had been obtained, whichever occurred first. Once five sessions at, or above, 90 % correct had been reached, threshold sessions began. All hens reached the criterion level of performance after 13 sessions.

Threshold sessions. The procedure for the threshold sessions was the same as for the successive presentation threshold sessions, except the flickering and steady lights were presented simultaneously. Each hen had a total of 10 successful threshold sessions in which the percentage correct in the first and last block of trials was at or above 90 % correct.

Results

Throughout this thesis, all raw data are presented in the Appendices located on the inside back cover. Only successful threshold sessions were included in these analyses. A total of 20 successive stimulus presentation threshold sessions (range, 1 to 6 for individual hens) and 22 simultaneous stimulus presentation threshold sessions (range, 0 to 11) were excluded. There were no systematic differences between the two presentation methods. Two of the 20 successive stimulus presentation sessions, and seven of the 22 simultaneous stimulus presentation sessions were excluded because the hens' performances did not reach the 90 % criterion in the first block of trials at the training flicker rate. All other sessions were excluded because the hens' performance failed to recover to levels above the 90 % criterion in the final block of trials at the training flicker rate. However, although these sessions have been excluded from analyses, they show very similar functions to the successful threshold sessions.

Figure 1.1 shows the percentage correct for all 10 successful threshold sessions for the successive stimulus presentation method. The functions show each flicker rate for all six hens. The filled in circle at the 25 Hz flicker rate represents the final block of trials at the training flicker rate. All hens showed very high accuracy (close to 100 %) at the lower flicker rates and a subsequent decrease in performance accuracy as the flicker rate was increased. The functions are extremely consistent across hens and are fairly consistent across threshold sessions. There is a tendency

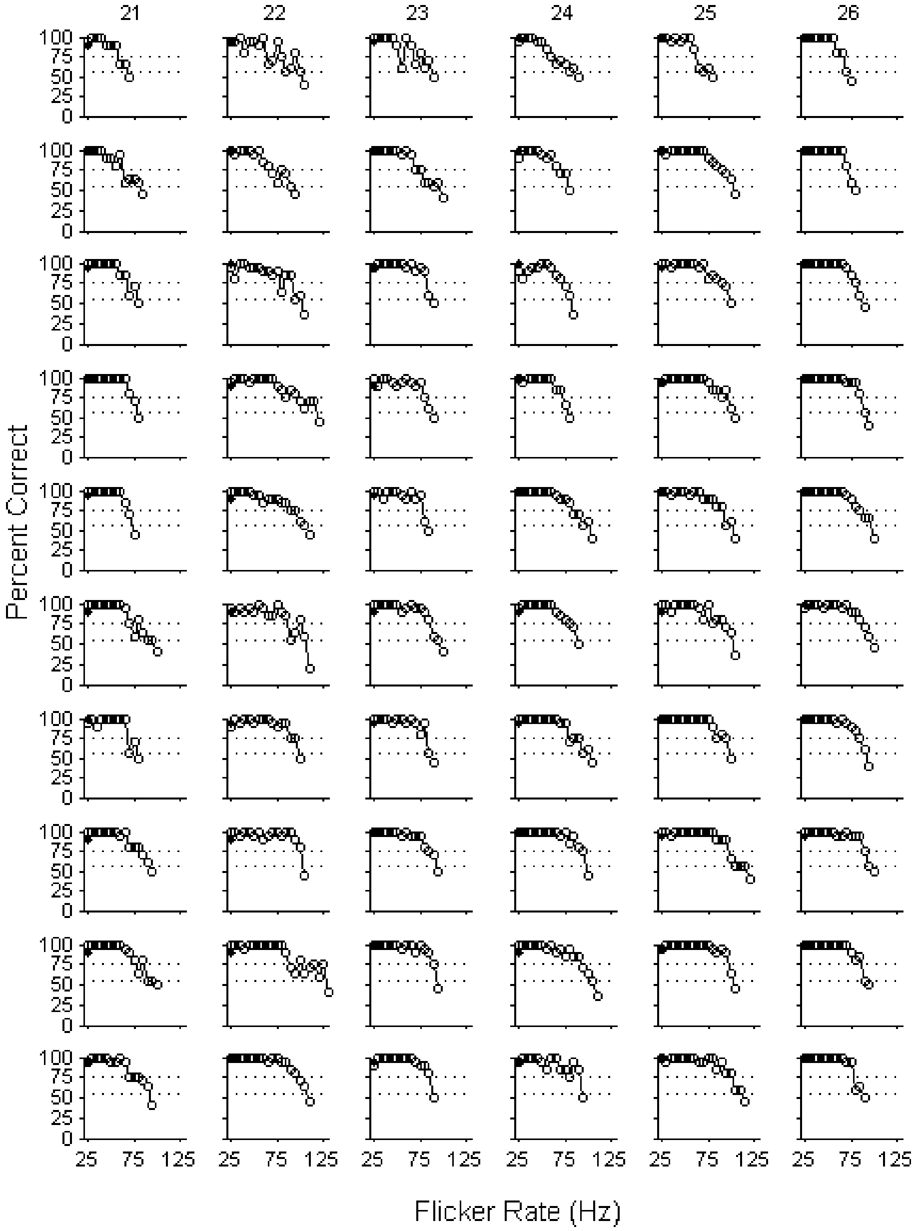


Figure 1.1. Percentage correct of all 10 successful threshold sessions for each hen for the successive stimulus presentation method. The functions show each flicker rate for all six hens. The filled in circle at the 25 Hz flicker rate represents the final block of trials at the training flicker rate. The horizontal dotted line marks 55 and 75 %.

for accuracy to be slightly better at some flicker rates in the later threshold sessions than in the early ones.

Figure 1.2 shows the percentage correct for all 10 successful threshold sessions for the simultaneous stimulus presentation method. The functions show each flicker rate for all six hens. The filled in circle at the 25 Hz flicker rate represents the final block of trials at the training flicker rate. As with Figure 1.1, the functions are consistent across hens and across threshold sessions. The functions are very similar to those shown in Figure 1.1, showing very little difference across the two methods of stimulus presentation.

Figure 1.3 shows the percentages correct, averaged (using means) over all 10 successful threshold sessions, on logarithmic Y-axes, for each flicker rate for all six hens. The left panel shows data from successive stimulus presentation and the right panel shows data from simultaneous stimulus presentation. As some higher flicker rates were reached on only a few threshold sessions, Figure 1.3 includes data only when there were three or more values contributing to the mean. The vertical lines mark one standard deviation each side of the mean. The horizontal dotted lines are at 75 and 55 % correct and the data points marked by a cross at 25 Hz (the training flicker speed) represent the percentage correct for the last block of trials of each threshold session. This cross is sometimes partially obscured. For all hens, and both stimulus presentation methods, percentages correct stayed high (close to 100 %) with little variability for flicker rates from 25-50 Hz. Performance decreased systematically for all hens beginning in the range from 55 to 75 Hz and falling to 50 % at flicker rates between 75 and 95 Hz. Variability increased for all hens as flicker rate increased and as percentage correct decreased. Figure 1.3 shows that there were no consistent differences across stimulus presentation methods.

The points where each hen’s averaged data line crossed 75% for the first time are presented in Table 1.1. The values are similar across methods of stimulus presentation and range between 69.8 and 95.4 Hz for successive presentation and between 68.5 and 92.3 for simultaneous presentation. A paired samples *t*-test comparing these values over methods was not statistically significant ($t(5) = 0.467$, $p > .05$). The largest difference in these values between the two stimulus presentations is 7.8 Hz for Hen 26. The values across hens are similar, except for Hen 21 where they are lower for both stimulus presentation methods. For any one

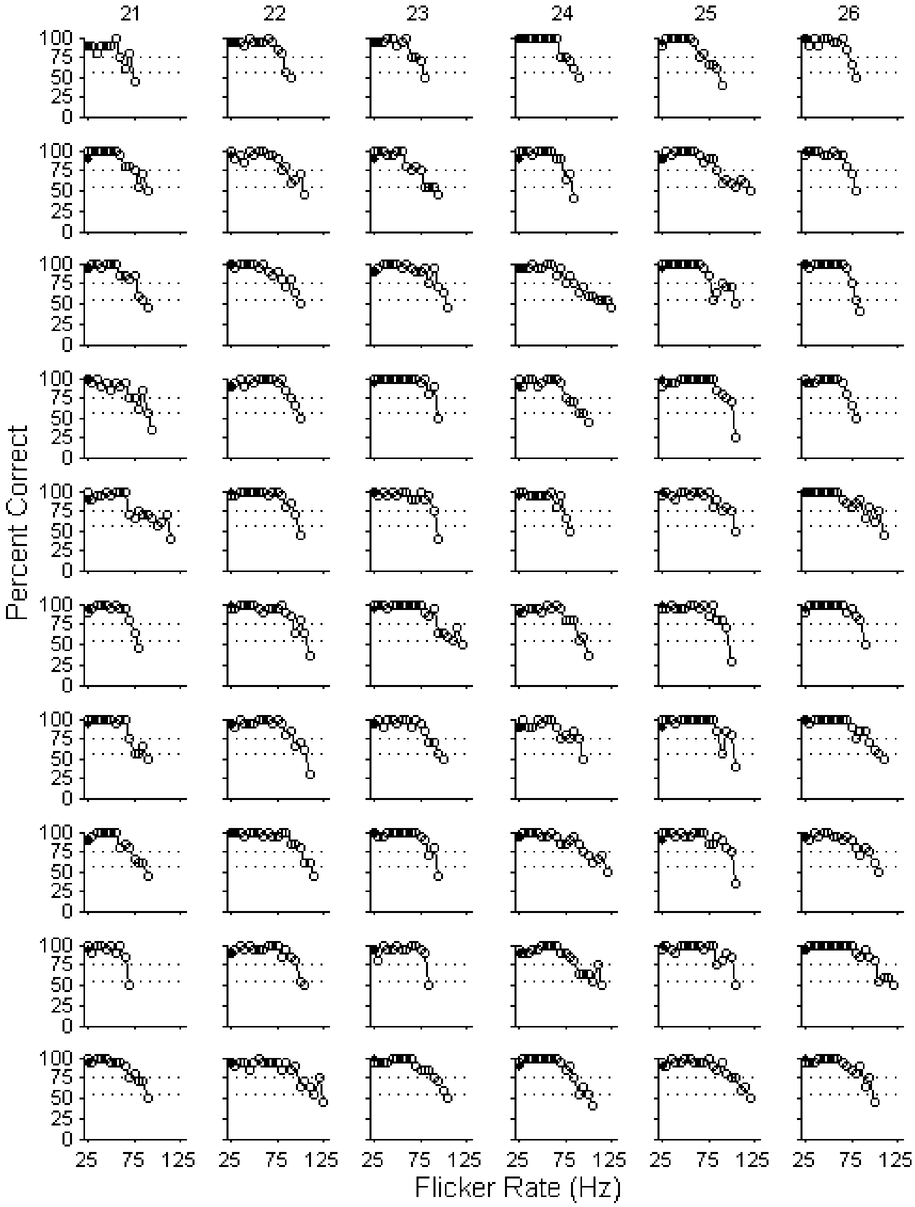


Figure 1.2. Percentage correct of all 10 successful threshold sessions for each hen for the simultaneous stimulus presentation method. The functions show each flicker rate for all six hens. The filled in circle at the 25 Hz flicker rate represents the final block of trials at the training flicker rate. The horizontal dotted line marks 55 and 75 %.

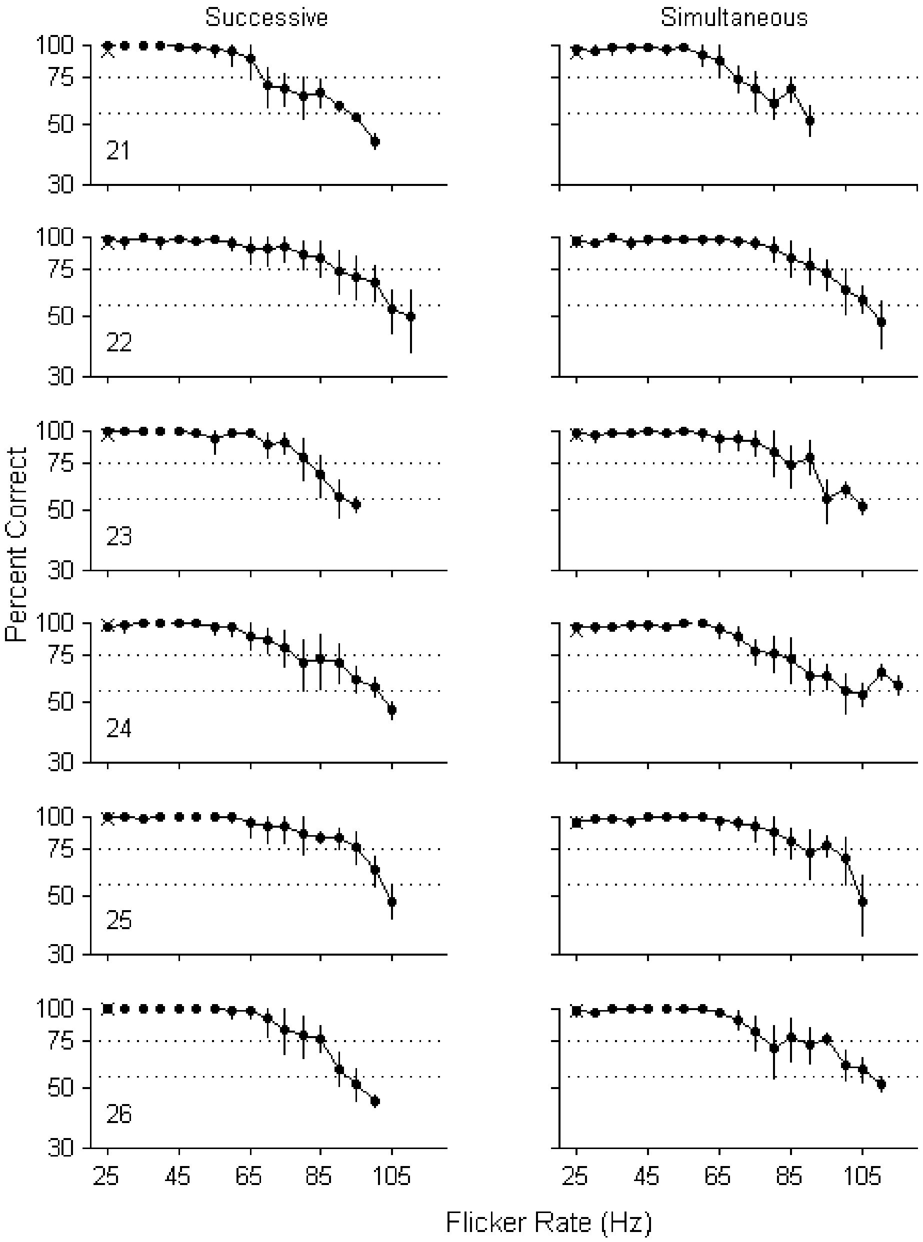


Figure 1.3. Percentage correct averaged over all 10 successful threshold sessions, on logarithmic y axes, for each flicker rate for all six hens. The left panel shows percentage correct for successive stimulus presentation and the right panel shows percentage correct for simultaneous stimulus presentation.

Table 1.1.

The point at which the mean percentage correct crosses 75 % for the successive and simultaneous stimulus presentation methods for each hen.

	Hertz values at 75 % correct	
Hen Number	Successive	Simultaneous
21	69.8	68.5
22	89.7	92.3
23	81.9	84.4
24	77.5	82.4
25	95.4	88.4
26	85.5	77.7
Mean	83.3	82.3

hen, the proportion of left responses in each threshold session did not deviate from 0.5 by more than 0.05.

The percentage correct for each flicker rate (70-105 Hz) across each of the 10 successful threshold sessions for the successive stimulus presentation method for all six hens are plotted in Figure 1.4. This was done to examine the effect of multiple threshold sessions in each 20 block trial. Figure 1.4 shows the functions for 70 Hz and above only, as accuracy below this for all hens was at, or close to, 100 %. A threshold session finished when responding fell below 55 % so there are missing data at the higher flicker rates indicating that rate was not reached in that threshold session. Hens 21, 23 and 26 did not complete threshold trials at 105 Hz, and so no data are shown at that flicker rate. The first data point comes from the first threshold session. It can be seen that for all hens at the slower frequencies, the first data point tends to be lower than those for further threshold sessions. In some cases, performance improved across threshold sessions. Since of interest is what flicker rate the animal can discriminate, any measures should probably exclude the earlier threshold sessions data. For most hens, their performance did not improve over the last five threshold sessions.

Figure 1.5 shows the percentage correct at each flicker rate (70-105 Hz) across each of the 10 successful threshold sessions for the simultaneous stimulus presentation method for all six hens. Again, hens showed a similar pattern to Figure 1.4, where the first data point was lower than the data points of the later threshold sessions for some frequencies. However, this appears at higher frequencies on Figure 1.5. The performance of Hens 21, 22 and 26 was more stable at the lower flicker frequencies than was their performance shown in Figure 1.4. This is possibly because the hens now had previous experience with the testing procedure.

Table 1.2 shows the median flicker rate at which each hen’s performance fell below 75 %, for the last 3 of the first 5 threshold sessions, and the last 3 of all 10 threshold sessions for both presentation methods. Generally, percentages correct for the earlier threshold sessions were lower than those of the final threshold sessions showing that the accuracy improved during the later threshold sessions for most hens.

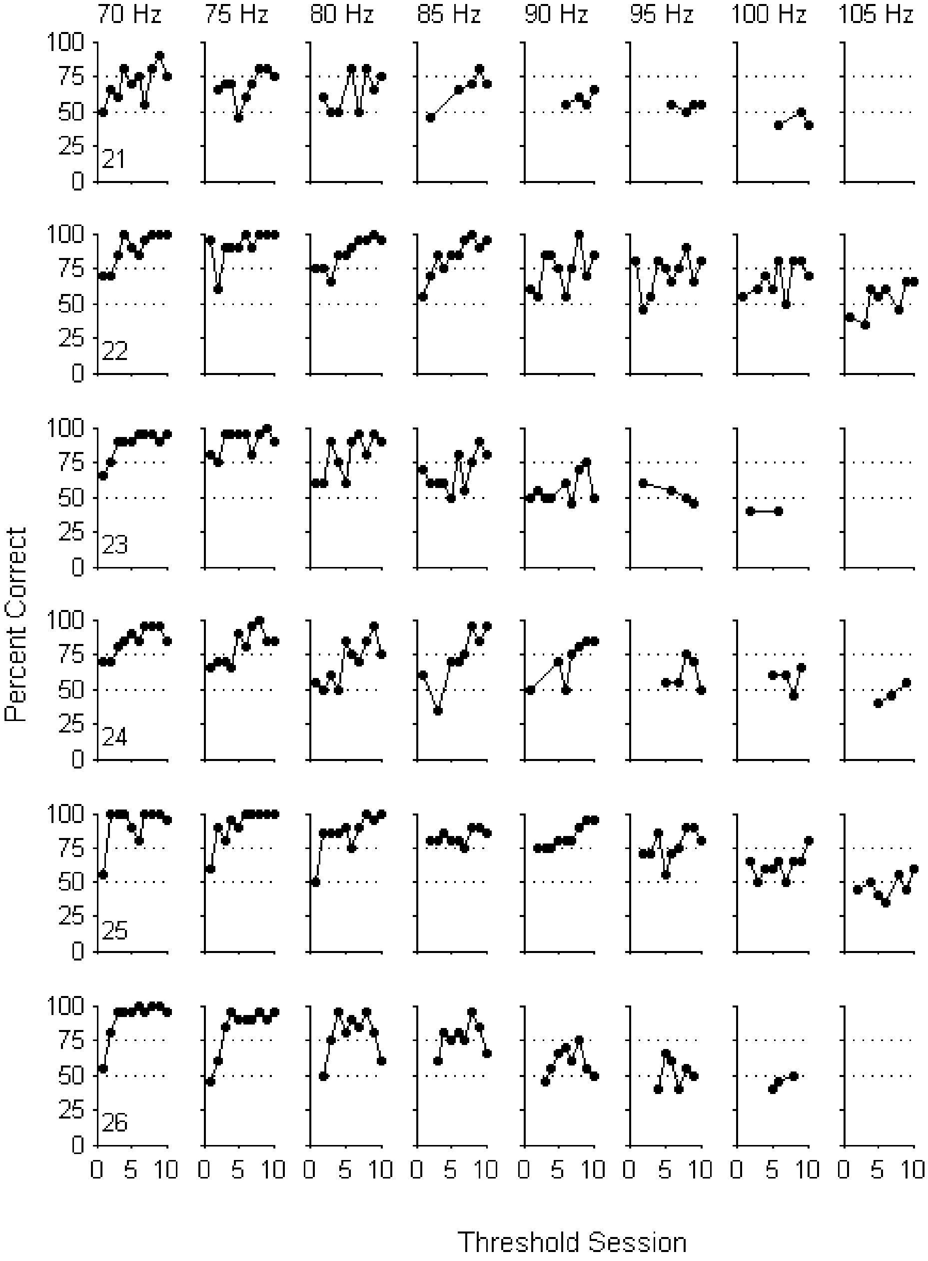


Figure 1.4. Percentage correct at each flicker rate (70-105 Hz) across each of the 10 successful threshold sessions for the successive stimulus presentation method for all six hens. The horizontal dotted lines mark 50 and 75 %.

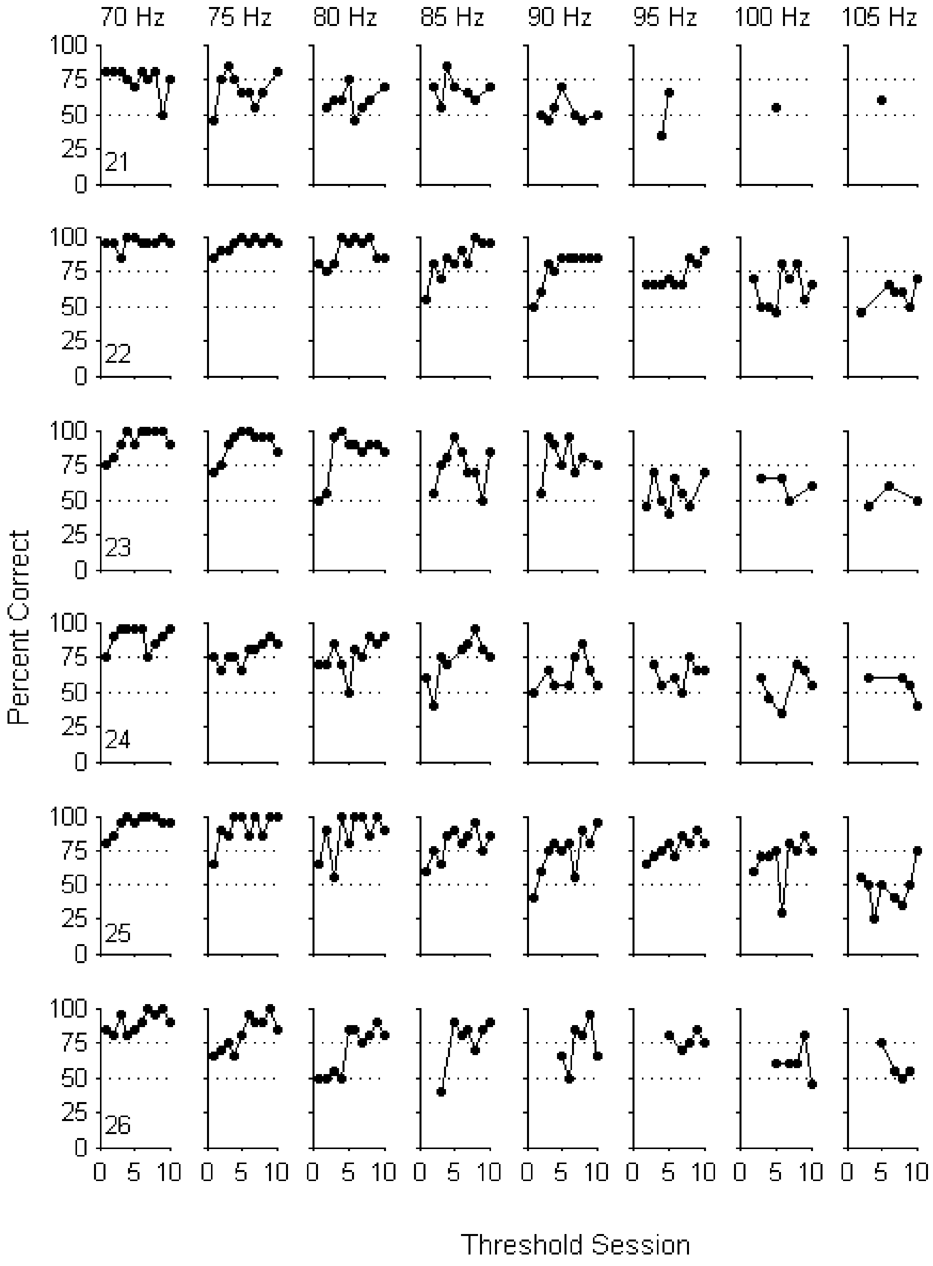


Figure 1.5. Percentage correct at each flicker rate (70-105 Hz) across each of the 10 successful threshold sessions for the simultaneous stimulus presentation method for all six hens. The horizontal dotted lines mark 50 and 75 %.

Table 1.2.

The median flicker rate that each hen's accuracy dropped below 75% correct for the last 3 of the first 5 threshold sessions, and the last 3 of all 10 threshold sessions for both successive and simultaneous stimulus presentation.

	Successive		Simultaneous	
Hen Number	Last 3 of 5	Last 3 of 10	Last 3 of 5	Last 3 of 10
21	75	90	80	75
22	100	100	95	100
23	85	90	95	85
24	75	95	80	90
25	95	100	100	105
26	90	90	80	100

Discussion

Both procedures in this study gave high percentages correct at low frequencies with systematically deteriorating performance at higher frequencies. Hence, both methods allowed estimation of a value for the CFF at the particular luminance used (300 cd/m^2).

For two alternative forced-choice (2AFC) and conditional discrimination procedures (such as those used in this study) the threshold is usually taken as the point where accuracy falls to 75 %. This is half-way between perfect responding and chance performance (Madigan & Williams, 1987; McKee, Klein & Teller, 1985). Table 1.1 gives the flicker rate corresponding to the 75 % point for the particular luminance used here. The CFF values were similar across hens, and both the simultaneous and successive procedures gave similar results. The CFF values found in the present study, using a white LED, ranged from 69.8 to 95.4 (successive presentation) and 68.5 to 92.3 (simultaneous presentation) which are similar to the values found by Nuboer et al. (1992) using a fluorescent lamp. These CFF values are much higher than the reported CFF of humans (approximately 50-60 Hz, Brundett, 1974; Hart, 1992) and thus provide support for the findings of both Nuboer et al. (1992) and Jarvis et al. (2002) that hens have higher flicker fusion values than humans.

The similarities of CFF found in both methods here suggest that the differences in CFF found by Jarvis et al. (2002) and Nuboer et al. (1992) were probably not a result of successive (Nuboer et al.) and simultaneous (Jarvis et al.) stimulus presentation methods. There are other differences between Nuboer et al. and Jarvis et al. including different light stimuli, different and occasionally unclear procedures, and different methods of analysis. It is likely that a combination of these may account for the differences found.

It is common practice in psychophysical research to do a small number of threshold sessions (commonly only one is used). However, the first few threshold sessions in this study, across both procedures, tended to underestimate the CFF thresholds. It may be that a number of threshold sessions are required in order to gain a more accurate picture of an organism's psychophysical threshold. In order to accurately assess CFF in animals, we need to establish the point at which they can no longer perceive the light as flickering. It may be that underestimation of the CFF thresholds of the first few threshold sessions are a result of the hens first learning to do the task. The hens have had an extensive number of trials at the training flicker rate (25 Hz), however, the threshold sessions increase this flicker rate away from the training stimulus. Therefore, the point at which they reach their threshold would tend to be lower as the threshold session is leading to a disruption in the previously learned task. More exposure to the threshold procedure allows the animals to learn the new task (i.e., flashing light vs. still light rather than flashing light at 25 Hz vs. still light), and is therefore a more stable measure of the limits of the animal's visual system – rather than a disruption in stimulus control as a result in the change of procedure. It may be that researchers should do a number of threshold sessions until responding becomes consistent. This would help to gain a more accurate representation of the psychophysical abilities of the animal being assessed.

There were a number of unsuccessful sessions in the present study that failed to reach criterion and were subsequently excluded from analyses. The unsuccessful threshold sessions were analysed and compared with the successful threshold sessions and it was found that there were no consistent differences. The stringent criterion used meant that only three incorrect trials in a final block of 20 would result in a percentage value of 85% correct and thus be deemed an unsuccessful session. As these final block of trials at the training flicker speed were unsignalled, hens may have continued to respond at chance levels until they were presented with a flickering trial when stimuli were presented successively. It would have been more appropriate to signal this return to training speed with a flicker trial and thereby avoid any possible carry over effects of random discrimination. However, this was only a problem when stimuli were presented successively and there were also a number of unsuccessful trials when stimuli were presented simultaneously. This would suggest the criterion of 90 % or above was too stringent during the blocks of trials with the training speed. Therefore, it was concluded that unsuccessful sessions were possibly a result of the criterion used rather than external effects such as fatigue, satiation, or the inability of the hens to do the task.

One implication of the finding that hens have high CFF thresholds is that they may have difficulty perceiving stimuli that are presented to them on conventional cathode ray tube (CRT) monitors. Patterson-Kane et al. (1997) hens' failure to learn to discriminate between two sets of moving video images may have resulted from the use of CRT monitors rather than an inability to learn the 2D discrimination. Hens may be able to perceive 2D images such as video if they do not flicker.

One possible way to overcome the problem of presenting 2D stimuli to species that have high CFFs may be to use a thin film transistor (TFT) monitor rather than the conventional CRT monitor. TFT monitors are virtually flickerless as there is no phosphor decay (the pixels do not darken) and so animals with high CFF thresholds may be better able to view stimuli presented on TFT monitors. For example, Ikebuchi and Okanoya (1999) found that Zebra and Bengalese finches emitted directed singing and showed courtship behaviour towards images of finches that were presented on a TFT screen but not when the same images were presented on a CRT screen. They concluded that the TFT images were responded to in a similar manner to conspecific birds, whereas CRT images were not.

In conclusion, the data from the present study show that the CFF thresholds established at one particular luminance level were comparable across the two procedures and that either is a useful method with which to determine CFF. In addition, the data show that hens have a higher CFF than that reported for humans. Therefore, any stimuli presented on standard CRT monitors may be seen as flickering and difficult for hens to see. Any future research using 2D images with hens should probably use TFT monitors rather than CRT monitors to avoid this problem. Further research is required to ascertain whether hens and other avian species will respond to TFT monitors when they fail to do so with CRT monitors.

EXPERIMENT 2

One method to overcome the problem of presenting 2D stimuli to species that have a high CFF may be to use thin film transistor (TFT) monitors rather than the conventional cathode ray tube (CRT) monitors. Computer and televisions screens have commonly been CRT screens, but there has been a recent move to using TFT screens.

As stated earlier, images on a CRT screen get refreshed at a particular rate. Each full screen scan refreshes the image with an electron gun which sweeps across the screen, lighting coloured phosphor dots to make up the image which darken between refresh scans. The higher the refresh rate of a monitor, the more frequently the image is updated. Images presented on screens with low refresh rates can appear to flicker as the phosphor dots fade before the next screen refresh.

One interpretation of the data from Experiment 1 is that, at 50-60 Hz, images may well appear to flicker for organisms with a CFF higher than this. Refresh rate is less important on TFT monitors than on CRT monitors, as each pixel remains lit between scans rather than turning on briefly and then turning off as happens in CRT monitors. As a result, there is no phosphor decay with TFT monitors and they are virtually flickerless. This suggests that CRT and TFT monitors may give rise to different findings. Although Chen and Lin (2004) found no difference in humans' ability to recognise the direction of a Landolt-C ring (either up, down, left or right) when they were presented on a CRT (which was set above the threshold for humans at 60 Hz) or a TFT-LCD (liquid crystal display) monitor, the different monitor types may affect animals' ability to recognise visual stimuli. In particular, animals with high CFF thresholds may be able to view and respond to stimuli presented on TFT monitors but may not be able to respond to stimuli on CRT monitors.

There are studies with birds that lend some support to this suggestion. Some authors have reported that birds failed to respond appropriately and/or failed to transfer a learned discrimination from CRT images to the real object (e.g., D'Eath & Dawkins, 1996; Patterson-Kane et al., 1997; Pepperberg, Naughton & Banta, 1998; Ryan & Lea, 1994). Patterson-Kane et al. (1997) found that hens did not generalise their discrimination of real objects to videoed images of the same objects presented on CRT monitors, and had more difficulty learning to discriminate between two sets of moving video images than between the real objects (except when the discrimination could be done on colour alone). Pepperberg et al. (1998) found grey parrots failed to learn new vocalizations from a videotaped CRT presentation of a conspecific model, although they learned successfully from live sessions with the same conspecific. D'Eath and Dawkins (1996) found that hens, exposed to either familiar or unfamiliar conspecifics behind Perspex screens, began to feed more quickly near a familiar hen; but did not respond in a similar manner when the condition was repeated with life-size colour CRT images of the same conspecifics. Ryan and Lea (1994) found that pigeons gave no social responses to a life-size moving video of a pigeon while they did with individual live pigeons. These studies suggest that the video images were not equivalent to the real stimuli for these birds. It is possible that the refresh rates of the monitors may have been a factor, however, the particular refresh rates of the monitors are not reported.

Studies with birds using TFT monitors have generally had more success in discriminating images than those using CRT screens. Ophir and Galef (2003) stated that TFT monitors may aid in producing more lifelike motion in visual stimuli and may be more effective than CRT monitors in "eliciting natural behaviour, especially in birds whose visual acuity, colour perception and high maximum critical flicker-fusion frequencies may make them particularly susceptible to the inadequacies of CRT displays" (p.370). TFT monitors have become more readily available and affordable, and are becoming more commonly used in behavioural research. Spetch and Friedman (2006) used a TFT monitor to present stimuli in a train-and-transfer procedure examining equivalence in pigeons. Two groups of pigeons were trained to discriminate between multiple views of two objects or between multiple views of pictures of the objects. In transfer tests, the discrimination generalised to images if they had been trained with objects, or to the objects if they had been trained with images. This finding showed that pigeons could respond to images on a TFT screen as they did to the real object. Female Zebra and Bengalese finches will show courtship behaviour to images of conspecific males presented on a TFT monitor, (Swaddle, McBride, & Malhotra, 2006) and female Japanese quail increased the time they spent near a live male after having seen the same male mate with another female in a video shown on a TFT monitor (Ophir & Galef, 2003) suggesting that in these studies the TFT images were equivalent to real conspecifics. Rieucau and Giraldeau (2009) used video playback on a TFT monitor to assess the effects of group size and companion type on the feeding behaviour of nutmeg mannikins. They found that feeding behaviour changed as group size changed similarly with both real and videoed companions, suggesting that video playback was an effective means to test social foraging. Further, Toda and Watanabe (2008) found that pigeons were able to discriminate between live streamed videos of themselves from pre-recorded video images presented on LCD monitors (which eliminate flicker similarly to TFT monitors). This shows that pigeons can view and discriminate images on LCD monitors. All of these studies suggest that video images may be equivalent to the real stimuli when images are presented on TFT screens.

In contrast, Pepperberg and Wilkes (2004) found that two Grey parrots failed to learn new vocalisations when tutored with video images presented on a LCD monitor. These results were similar to those found by Pepperberg, Naughton and Banta (1998) using CRT screens. However, Pepperberg and Wilkes (2004) conclude that it was the lack of social interaction in general, rather than the screen used (whether CRT or LCD), that inhibited the learning of new vocalisations by the parrots.

If it is the refresh rates of CRT screens that gave rise to the lack of generalisation found in some studies, it seems reasonable to suggest that increasing the refresh rate above the CFF of the animal should lead to results similar to those found with TFT screens. A study that provides some support for this was conducted by Galoch and Bischof (2006). They examined zebra finches' discrimination between two live video images presented on a CRT monitor set at 100 Hz. This is a higher refresh rate than conventional CRT monitors can generally be set. The birds spent more time near images of a zebra finch rather than an empty cage, and images of an unknown female rather than an unknown male. The authors conclude that the birds could recognise and respond to the video images. Further support comes from a study by D'Eath and Dawkins (1996). These researchers found that hens took longer to feed near a bird showing a threat-like posture in a video presented on a CRT monitor set at 100 Hz. The hens also behaved differently to videos that contained hens and those that didn't, implying that they were able to recognise the images and respond appropriately. However, other parts of this study suggest that not all video images gave comparable results to real stimuli. For example, images of unfamiliar hens had less effect than the unfamiliar hens themselves. Despite this, the findings of these two studies suggest that increasing the refresh rate of a CRT monitor may aid the birds' ability to discriminate or respond to stimuli presented on such monitors.

The results of the studies outlined above suggest that there may be differences in a hen's response to stimuli presented on a CRT and a TFT monitor. One way to assess this would be to test whether hens are able to show correspondence between two images of the same object/stimulus. One method would be to present images of stimuli which would normally produce specific spontaneous responses (e.g., food, conspecifics or prey species) on a TFT and CRT monitor and assess whether animals respond in a similar manner to both monitors. Ikebuchi and Okanoya (1999) used this method to compare finches' responses to CRT and TFT monitors. They found that male Zebra and Bengalese finches emitted directed singing, and showed courtship behaviour, towards images of conspecific females presented on a TFT screen comparable to that found with live conspecific females. The finches failed to do so when the same images were presented on a CRT screen set at 60 Hz. These results imply that the TFT image did appear the same as the real conspecifics, while the CRT image did not.

Another method to assess whether hens respond to both types of screens in a similar manner would be to train hens, using a conditional discrimination procedure, to discriminate between two stimuli presented on a TFT monitor, and test for transfer of that learned response to the same stimuli on a CRT monitor. To the author's knowledge, there are no studies that have used a learned response to examine if birds respond in a comparable manner between two screen types.

Conditional discrimination procedures are a commonly used method to train animals to discriminate stimuli. In a conditional discrimination, two or more stimuli are presented and the animal is required to make a response to two or more manipulanda. The correct response depends on which sample stimulus was presented. For example, in a two stimuli conditional discrimination, the animal is required to respond to one key in the presence of one stimulus and respond to the other key in the presence of the other stimulus, which Sidman and Tailby (1982) described as an "if…then" relation. Past research with hens has shown that they can learn conditional discriminations (e.g., DeMello, Foster, & Temple, 1992; Temple, Foster & O'Donnell, 1984). Therefore, the present study used a conditional discrimination procedure to train hens to discriminate between two stimuli on a TFT monitor, and then looked for transfer to the CRT monitor.

Once the hens had learned to discriminate the stimuli on the TFT monitors, transfer tests to the CRT monitors were used to give an indication of how the different monitor types affected performance. The transfer tests involved changing the screen that the stimuli were presented on. If performance remained high during the transfer tests, this showed generalisation across screen types. However, if performance decreased, then perception of pictorial stimuli may have been affected by the type of screen used. Transfer or generalisation tests often involve presenting the alternative stimuli for a percentage of trials throughout the normal training procedure. For example, Jitsumori, Sieman, Lehr and Delius (2002) assessed equivalence classes in pigeons where 32 of the 160 trials (20 %) were test trials, and Zentall, Friedrich, and Clement (2006) examined timing in pigeons where 56 of the 200 trials (28 %) were test trials. However, as the transfer tests in the current experiment required changing the computer screen used to present the stimuli, transfer tests ran for a whole session, rather than a number of trials.

Many studies using a conditional discrimination procedure include an observing response before the presentation of stimuli. An observing response is a response that must be completed before the stimulus is turned on, but that has no effect on the probability of reinforcement (Wyckoff, 1952; Zeigler & Wyckoff, 1961). For example, in a study by Zeigler and Wyckoff (1961) pigeons were shown the discriminative stimuli only after depressing a pedal that was located in front of the response keys. More often, observing responses consist of a fixed number of responses to a key or screen before exposure to the discriminative stimuli (e.g., Alsop & Jones, 2008; Davison & McCarthy, 1989; DeMello et al., 1992; Emmerton & Renner, 2006; Friedrich & Zentall, 2010). The advantage of using an observing response is that they orient the animal towards the discriminative stimuli, and Zeigler and Wyckoff (1961) state that it increases the probability that the animal is attending to the stimuli when responding. As such, Zeigler and Wyckoff state that observing responses "play an important role in the acquisition of discrimination" (p. 131).

While observing responses are common in studies using a conditional discrimination procedure, there are some studies that have found successful discrimination without including an observing response (e.g., Signal, Temple & Foster, 2001; Temple, Foster & O'Donnell, 1984). In the study by Temple, Foster and O'Donnell, hens learned an auditory conditional discrimination task without an observing response, showing that hens were able to attend to stimuli without requiring an observing response to orient them. Observing responses work much the same in auditory procedures as they do in visual procedures in that they signal the start of a trial and help to ensure the hens are in a position to see or hear the stimuli.

The present experiment involved two keys presented on each side of a computer screen on which the images were presented. An appropriate observing response would be to peck a key present on the screen directly below where the images were to be presented. While the TFT monitor used during training was a touch screen and could have been used to record pecks, this was to be replaced during test sessions with a CRT monitor that could not record pecks. Thus, as hens had been shown to successfully learn a conditional discrimination without an observing response (see Temple, Foster & O'Donnell, 1984), and in light of the technical limitations of the CRT screen, it was decided not to include an observing response in the current experiment to keep the training and testing procedures as similar as possible.

The main aim of this experiment was to establish a procedure to assess whether images presented on different screens are perceived as equivalent by hens. The first phase of this experiment aimed to investigate whether hens showed similar responding to a TFT screen and a CRT screen set below their CFF. Hens were trained to discriminate between two coloured stimuli (red and green squares) presented on the TFT monitor. Colour was used as hens are able to discriminate easily between colours and many studies have used colour as a discriminative stimulus. For example, Jones, Carmichael and Rayner (2000) found that chicks preferred white or yellow string pecking devices to red, green or blue ones. Dawkins and Woodington (1997) trained chickens to discriminate between blue and red objects. Huber-Eicher (2004) and Zupan, Kruschwitz, and Huber-Eicher (2007) found hens laid more eggs in yellow coloured nest boxes, over blue, green or red nest boxes, showing that the hens could distinguish between the colours. Foster, Temple, Mackenzie, DeMello and Poling (1995) found hens were easily able to distinguish between red and green keys that were used as sample and comparison stimuli in a matching-to-sample task. The first phase of the present experiment, then, involved a conditional discrimination in which a hen was required to peck a left key if the stimulus red and the right key if the stimulus was green.

There are several ways of evaluating performance under such a procedure. One common measure is percentage correct over all trials. However, this does not show any biases the animal might have to selecting the left or right keys over and above the effects of the trained discrimination. To explore such biases further requires examination of percentage correct on the two different trial types and comparison of their data paths. This was done here.

Davison and Tustin (1978) suggested behavioural detection theory could provide a measure of an animal's ability to discriminate between stimuli not confounded by response biases. They proposed log *d* as a bias-free measure of the degree to which an animal discriminates between stimuli. The equation for calculating log *d* is:

$$Log\ d = 0.5(log(W/X) - log(Y/Z)) \qquad (1)$$

where W is the number of correct responses to the stimulus associated with the left key, X is the number of incorrect responses to the stimulus associated with the left key, Y is the number of correct responses to the stimulus associated with the right key and Z is the number of correct responses to the stimulus associated with the right key. A high ratio of correct (W and Y) responses to incorrect (X and Z) responses show high discriminability of the stimuli. Thus, as discriminability increases, so does log *d* (Davison & Tustin, 1978). Davison and Tustin (1978) also proposed log *c* as a measure of response bias. The equation they proposed for calculating log *c* is:

$$Log\ c = 0.5(log(W/X) + (log(Y/Z)) \qquad (2)$$

where W, X, Y and Z are as in Equation 1. Both log *d* and log *c* were calculated for the data in the present study and these were compared to the percentage correct measures.

Overall percentage correct was used to assess ongoing performance in the present experiment. A hen was deemed to have learned the discrimination if her accuracy was at or above 85 % correct over five, not necessarily consecutive, sessions. This accuracy level was selected because it was significantly above chance responding. That is, if each trial is regarded as a binomial "coin toss" and if there are 100 trials per session, any percentage correct over 59 % is significantly different from chance at a .05 level of significance, so if the hens' performances were at or above 85 % correct, they were responding well above chance. Once at least two hens were responding at or above 85 % correct, there were test sessions with the same stimuli presented on a CRT monitor.

The second, third and fourth phases of this experiment used stimuli that could be discriminated on the basis on shape rather than colour (a black cross and a black circle). A number of the studies outlined above used conspecifics as stimuli. The problem with using such stimuli is that they can often differ on a number of features. Therefore, transfer can occur on a number of features such as colour, movement, or overall size. To avoid this possibility and ensure that any other cues were ruled out, the stimuli used in the present experiment were a black cross and a black circle presented against a white background. The stimuli were such that they were the same shape and colour, and thus, differed in shape only. However, there were problems in gaining appropriate stimulus control during this phase. It was considered possible that the hens' failure to learn the discrimination during this phase may have been due to the stimuli being too large. It could have been that the stimuli extended too far beyond the hens' binocular visual field to be easily discriminable. Eye movement in hens is limited, and while they have a field of view that is approximately 300 degrees, they only have a 25-30 degree overlap in which binocular vision can occur (Prescott, Jarvis & Wathes, 2004). Hens have an accommodation of 17 diopers (D) (Schaeffel, Howland & Farkas, 1986) and lower field myopia, which allows objects in this field to be focussed on the retina at small viewing distances (Dawkins, 2002). In a study assessing the effect of stimulus size on object recognition, Lombardi and Delius (1990) trained pigeons to select the nonmatching stimulus in a nonmatching-to-sample task using stimuli of the same size, and then tested for generalisation to different sized stimuli. They found that the pigeons were able to generalise their discriminative performance to the differently sized stimuli, and that they were more accurate when the comparison stimuli were smaller than the sample stimuli. Other studies have found that pigeons generalised their performances in a same-different conditional discrimination (Peissig et al., 2006), and four-alternative-forced-choice procedure (Castro & Wasserman, 2010), to stimuli of different sizes (both smaller and larger). As a result of the finding from Lombardi and Delius, the sizes of the stimuli used in the present experiment were reduced. As these attempts to achieve control failed, there were no transfer tests in this phase.

In the third phase of this experiment, in an attempt to gain stimulus control, the training procedure was changed and an observing response was included using the TFT monitor. As pecks to the CRT monitor could not be recorded, the testing procedure had to be different from the training procedure as it was not possible to record the occurrence of an observing response, and so the stimuli had to be presented after a set period of time had passed. Transfer tests to the same stimuli presented on a CRT monitor set at 60 and 75 Hz were conducted.

The fourth phase was the same as Phase 3, except that transfer tests used a CRT monitor set at 100 Hz.

Method

Subjects

The subjects were the same as those used in Experiment 1 except that in this experiment Hen 21 died of skin cancer after Condition 1 and was replaced by another hen also identified as Hen 21 for all following conditions. The housing conditions were the same as Experiment 1 except that the individual cages in which the hens were housed were increased to 500-mm long × 510-mm wide × 420-mm high.

Apparatus

The experimental chamber (410-mm high × 580-mm long × 540-mm wide measured internally) was made of 20-mm thick black particle-board. The floor of the chamber was covered with removable grey artificial grass matting (Astrograss^{TM}). Two white 28 v back-lit circular response keys (30-mm diameter) were positioned on the right hand wall of the chamber 400 mm off the floor and 300 mm apart. Each key was surrounded by a metal plate (70-mm wide × 140-mm high). A brief feedback beep sounded when a response with a force of 0.1 N was made on either lit key.

Two apertures (70-mm wide × 100-mm high) located on each side of the experimental chamber allowed access to two magazines containing wheat. When a magazine was operated, the hopper was raised and a 1-W white light bulb illuminated the magazine for a 3-s period.

Between the keys a section of the chamber wall (200-mm wide × 215-mm high) was removed and a short (115-mm deep) open box was attached. This allowed placement of either a 15 inch TFT monitor (model 710A) or a 21 inch CRT monitor (Trinitron Multiscan) in such a way that the hens could see only the central area of the screens. Two computers were used throughout the study. One computer (Optiplex GX110) controlled the stimuli presentation, and a different computer (Optiplex GXa) controlled the experimental equipment and recorded all data using Med-PC® software (Version 4). The luminance of both screens were similar to the luminance of the key lights used in Experiment 1 (300 cd/m^2). Total session data were also manually recorded into a data book at the end of each session.

Procedure

Discrimination training.

Pre-training. The hens had been previously trained to peck a lit key in Experiment 1. The sample stimuli presented on the screens were a red or a green square (108-mm wide × 108-mm high) shown against a black background (see Figure 2.1). The stimuli were presented in the centre of the computer screen. For the first five trials, only the correct key was lit for each trial type. A trial began when one of the stimuli was presented on screen. After 1 s, the key associated with that stimulus was lit; left key on a red stimulus trial and right key on a green stimulus trial. The red or green stimulus remained on screen until a peck to the lit key operated the magazine associated with that key (left magazine with left key and right magazine with right key) and turned both the stimulus and key light off, ending that trial. After these initial five training sessions, both key lights were lit in further training sessions, but only pecks to the key associated with the stimulus presented were reinforced.

Figure 2.1. Red and green stimuli used in Condition 1.

General Procedure, Conditions 1-4. Each trial began with one of the stimuli being presented on the TFT screen for 1 s, after which both key lights were lit. A correct response resulted in the key lights being turned off and a 1-W white light-bulb illuminating the left magazine on one trial type and the right magazine on the other trial type for 3 s. If a reinforcer was scheduled to occur, the appropriate magazine was raised for 3 s. An incorrect response resulted in both key lights being turned off for a 3-s period (blackout). A trial ended after the magazine light (and reinforcer if scheduled), or a blackout period, had occurred. The next stimulus presentation signalled the start of a new trial.

Table 2.1 presents a list of the conditions, the stimuli, number of sessions that the VR was increased, final VR requirement, number of sessions in each condition and number of hens responding at or above 85 % for at least five sessions. The presentation of a stimulus was controlled pseudo-randomly according to a predetermined series to ensure that the same stimulus was not presented more than three times sequentially, and the total number of each trial type would be approximately equal within a session.

Reinforcement was scheduled according to a variable ratio (VR) schedule, in which reinforcers were programmed for a correct response after a variable number of correct responses. This VR requirement was increased over a number of sessions. Decisions to increase the VR were based on a combination of relatively stable performance and the subject having received approximately 30 reinforcers during the experimental session. The maximum VR varied over conditions (see Table 2.1). When a reinforcer was not scheduled for a trial, a correct response resulted in 3-s illumination of the magazine light, but the magazine was not raised.

In addition to the VR requirement, reinforcer delivery was controlled so that the rate of reinforcement was equal across any two trial types. Reinforcers were allocated to a trial type pseudo-randomly. For example, if the next reinforcer was scheduled for a correct response on one trial type, any correct responses on the other trial type were followed by the magazine light, but no food presentation, until the scheduled reinforcer for a correct response on the first trial type had been collected. Correct responses of both types were counted towards the VR requirement of the pre-selected trial type and a reinforcer was delivered for the first correct response that completed that VR. By controlling the rate of reinforcement in this manner, the VR requirement was sometimes increased slightly, particularly if the hen responded mainly to one key.

Sessions were terminated after 40 min had elapsed or after 30 reinforcers had been obtained, whichever occurred first. Once at least two hens' accuracy was at or above 85 % (and little to no biases were shown), the test sessions began.

Table 2.1.

The order of experimental conditions, together with the stimuli, final VR, number of sessions to VR, number of sessions in each condition, and number of hens responding at or above 85 % for at least five sessions, and whether there was an observing response.

Condition	Stimuli associated with left key	Stimuli associated with right key	Final VR	Sessions to VR	No. of sessions	No. of hens responding at or above 85 %	Observing response
1	Red	Green	10	27	136	6	N
2	Cross	Circle	10	33	70	4	N
3	Circle	Cross	8	29	44	0	N
4	Small circle	Small cross	7	31	54	0	N
5	Small circle	Small cross	5	27	60	6	Y
6	Small circle	Small cross	5	-	75	6	Y

The stimuli used in Condition 1 were a red or a green square (108-mm wide × 108-mm high) shown against a black background (see Figure 2.1). In Conditions 2 and 3, the stimuli were a black cross (plus sign) or a black circle (measuring 60-mm wide × 60-mm high) set against a white background, as shown in Figure 2.2. The same cross and circle stimuli were used in Conditions 4 to 6, but measured 21-mm wide × 21-mm high.

Figure 2.2. Cross and circle stimuli used in Conditions 2 to 6.

Test sessions with coloured stimuli (Condition 1). Test sessions were the same as the General Procedure, except that the stimuli (red or green square) were shown on the CRT monitor set at a refresh rate of 60 Hz rather than on the TFT screen. Each hen completed three consecutive test sessions, after which the hens continued on the training procedure until the commencement of Condition 2.

General Procedure, Conditions 5 and 6. After the failure for the hens accuracy to reach 85 % or above in Conditions, 2, 3 and 4, the procedure was changed. A red square (measuring 23-mm wide × 23-mm high) was presented on the screen below each stimulus. As the TFT was a touch screen, any pecks to the red square were recorded and a Fixed Ratio (FR) requirement was put in place. Under the FR, the hens were trained to peck the red square on the TFT monitor a fixed number of times before the two key lights were lit and operable, to ensure the hens were oriented toward the stimulus before pecking a key. In Condition 5, the FR requirement was increased over 11 sessions from FR 1 to FR 5. The VR requirement was lower in Conditions 5 and 6 (see Table 2.1), than in the previous conditions, as each trial took longer to complete. Training in Condition 6 started at VR 5, as the hens had previous experience with the procedure and stimuli. Test sessions occurred after Conditions 1, 5 and 6 as these were the only conditions where the hens' accuracy was at or above 85 % correct.

Test sessions with geometric shapes (Conditions 5 and 6). There were two types of test sessions; one that used a TFT screen and one that used a CRT screen. The TFT test sessions presented the stimuli on the TFT monitor used during training, and the CRT test sessions presented the stimuli on a CRT monitor set at different refresh rates. There were four test sessions in total for each condition. Table 2.2 presents a list of the order of test sessions, the screen type, and screen refresh rate.

As pecks to a CRT monitor could not be recorded (when they could be on the TFT screen), the FR observing response requirement was not possible during test sessions. The average amount of time it had taken hens to fulfil the FR requirement of the observing response in the training sessions was calculated and found to be 4 s. As a result, during all test sessions (both TFT and CRT), the stimuli were still presented with the red square directly below (as in the training sessions), but was automatically removed after 4 s had elapsed, regardless of any response from the hen. When the red square was removed, both key lights were lit. The stimulus remained on screen, and the key lights remained lit, until the hen pecked one of the keys. A correct response (a peck to the left key when the stimulus was the circle, and a peck to the right key when the stimulus was the cross) resulted in the stimulus being removed from the screen, both keys lights being turned off, and the appropriate magazine being illuminated for 3 s. In addition, the magazine was raised for 3 s if a reinforcer was scheduled to occur. An incorrect response resulted in the response keys being turned off and a 3-s black-out period before the next trial began. There were at least two training sessions between every test session.

Table 2.2.

The condition number, order of test sessions, the monitor type, and monitor refresh rate.

Condition Number	Test Session	Monitor Type	Refresh Rate
5	1	TFT	-
5	2	CRT	60
5	3	TFT	-
5	4	CRT	75
6	5	CRT	100
6	6	TFT	-
6	7	CRT	100
6	8	TFT	-

Results

Occasionally, the hens laid eggs during experimental sessions resulting in atypical performance. Thus, throughout this thesis, data were discarded if an egg was laid during the session. The axes scales for all graphs were kept constant so that direct comparisons could be made across graphs.

For all statistical tests, effect sizes were also calculated. For repeated-measures ANOVA, partial eta squared (η^2) are reported, and for all paired-samples *t*-tests, Cohen's *d* are reported as suggested by Aron, Aron and Coups (2009). Ferguson (2009) points out that null-hypothesis significance testing is limited by sample size. As only six subjects were used in the present experiment, effect sizes may give a clearer indication of the differences in accuracy across test sessions. Thus, as recommended by Cohen (1994) and Ferguson (2009), effect sizes are reported alongside the *p* values for the statistical analyses throughout this thesis. Ferguson (2009) suggests that for η^2, values larger than .41 shows a recommended minimum effect size recommending a "practically" significant effect for social science data (RMPE), and values above 1.15 and 2.70 show moderate and strong effects respectively. In addition, Ferguson suggests that for Cohen's *d*, any values larger than .04 shows a minimum practical effect size, and values larger than .25 and .64 show moderate and strong effect sizes respectively. Thus, these values were used throughout this thesis to determine the strength of any effect sizes. Theil tests (Hollander & Wolfe, 1973) were also conducted. Theil tests are non-parametric tests of trend and show whether the trend in the data for a hen is statistically significant (Hollander & Wolfe, 1973).

The left panel of Figures 2.3a shows the overall percentage correct and that of 2.3b shows log *d* over all training sessions for Condition 1 (the red and green stimuli). As there were some sessions where hens reached 100% accuracy, and as a result log values could not be calculated, the Hautus (1995) correction was used in all log calculations. This involves adding 0.5 to all the numbers of responses of each type and has been described as the most appropriate way to deal with these types of calculations (Hautus, 1995). All hens' accuracy was above 85 % correct relatively quickly (marked by the vertical dotted line), taking between 16 and 30 training sessions. In all remaining training sessions their performances remained consistently above 85 % correct before and after the test sessions (the locations of test sessions are shown by asterisks on the x axis, but the data are not plotted) showing that the

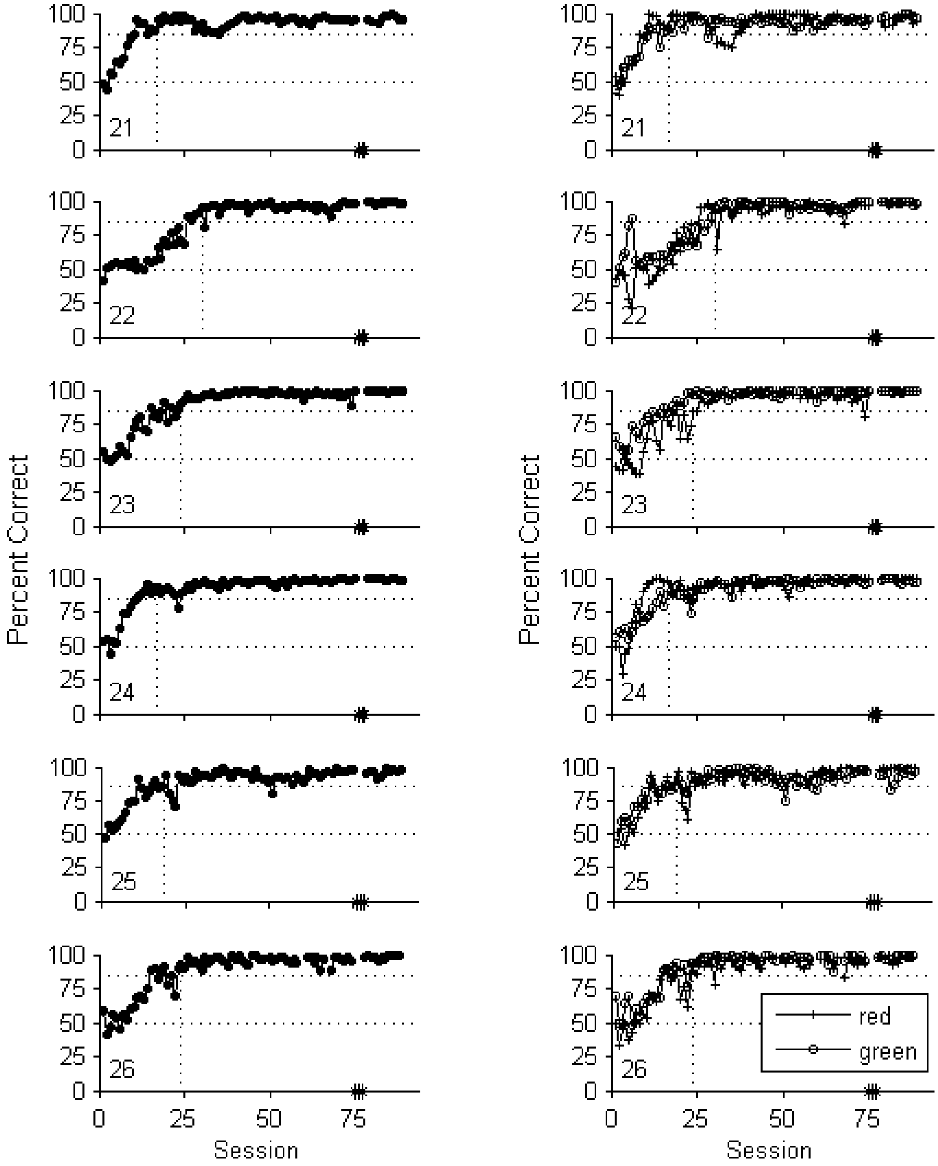


Figure 2.3a. The left panel shows the overall percentage correct over all training sessions for the red and green stimuli (Condition 1) for all hens. The right panel shows the percentage correct for each of the stimuli over training sessions. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 % and 85 % accuracy. The asterisks on the x axis represent the locations of the test sessions.

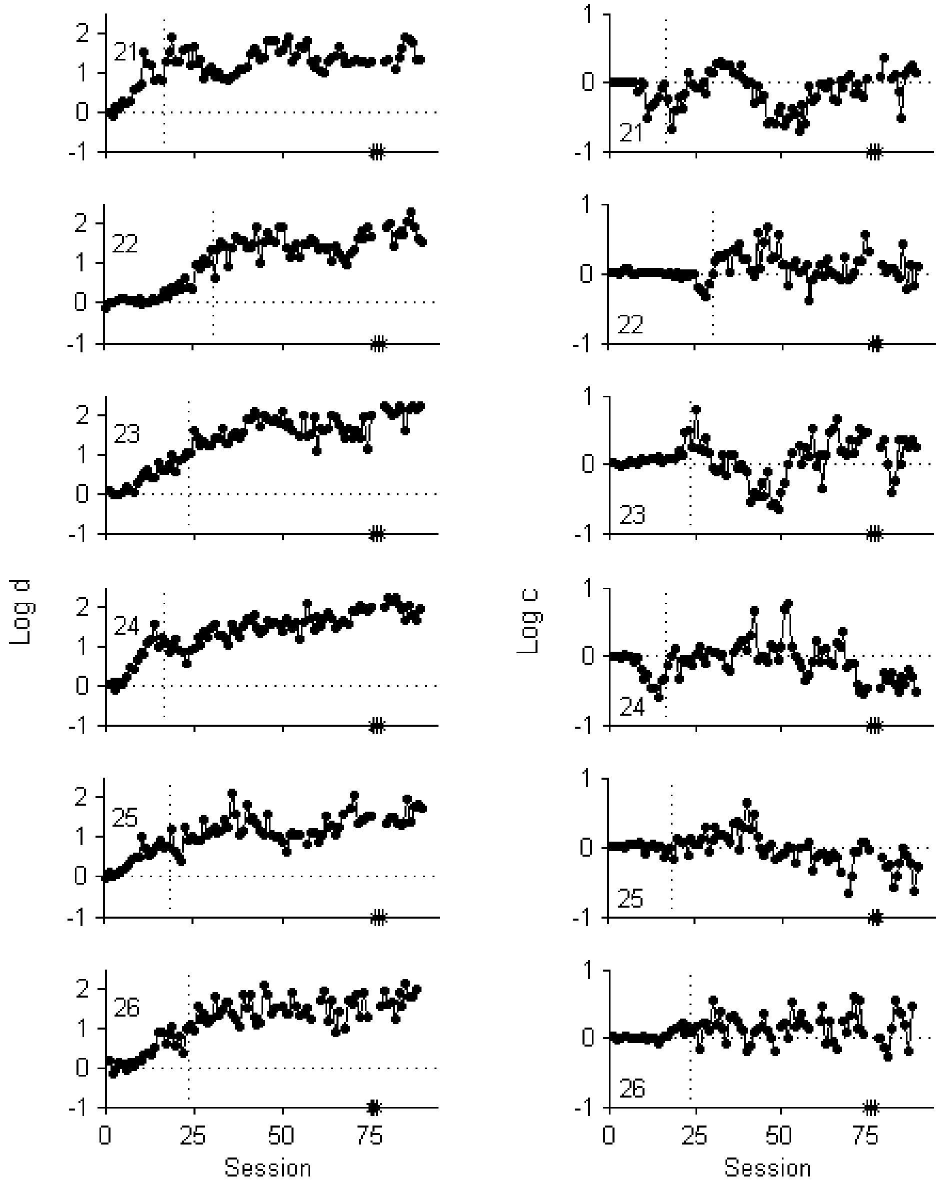


Figure 2.3b. The left panel shows the log *d* plots over all training sessions for the red and green stimuli (Condition 1) for all hens. The right panel shows the log *c* plots over training sessions. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

test sessions did not disrupt performance. The log *d* estimates also show that all hens had high discrimination. A CRT monitor that could go up to 100 Hz was not available at the beginning of this experiment, and as a result, each hen completed 75 training sessions while a monitor was sourced before the first test session.

On the right panel of Figures 2.3a and 2.3b, the percentages correct for each stimulus (red and green) and log *c* estimates across sessions for Condition 1 are shown. There were no consistent differences in accuracy of responding across the two stimuli with the percent correct plots, with all hens showing high accuracy on both red and green trials over the later sessions. Log *c* values that fall below 0 indicate a bias towards responding on the left key and log *c* values above 0 indicate a bias to the right key. There were no consistent differences in log *c* across hens over the later training sessions, with two of the hens showing no biases (Hens 21 and 22), two hens responding more to the left key (Hens 24 and 25), and two hens responding more to the right key (Hens 23 and 26).

Figure 2.4 shows the mean percentage correct and standard deviations for the last five training sessions for Condition 1 for all hens. Also plotted are the percentages correct for the three test sessions for all six hens. For all hens, the data from the last five training sessions show very high accuracy (range, 89 to 99% correct) and little variability. Accuracy for all test sessions remained high for all six hens (range, 88 to 99 % correct) and a repeated-measures ANOVA showed no significant difference across the training or test sessions ($F(3,15) = 0.400$, $p>.05$, $\eta^2=0.074$, which is below the RMPE suggested by Ferguson, 2009). There were no significant data trends for any of the hens (Theil test, C = 2, -3, -, 2 -5, and -2, for Hens 21 to 26 respectively, n= 4, p>.05 in all cases).

Figures 2.5a and 2.5b shows the overall percentage correct and log *d* plotted against all training sessions for Conditions 2, 3 and 4. Log *d* and percentage correct follow the same pattern. During Condition 2, accuracy for four hens (21, 23, 25 and 26) was at, or above 85 % correct within 25 to 43 sessions. Accuracy for Hens 22 and 24 was never above 85% correct for five sessions. The keys associated with each of the stimuli were reversed in Condition 3 to investigate a potential side/stimulus bias. It can be seen that no hen showed accuracy above 85 % correct over the 44 training sessions, although Hens 23 and 26 showed accuracy above 50 % correct and log *d* estimates that were trending upwards over sessions. The hens received at least 280 trials within a session for the last five training sessions in

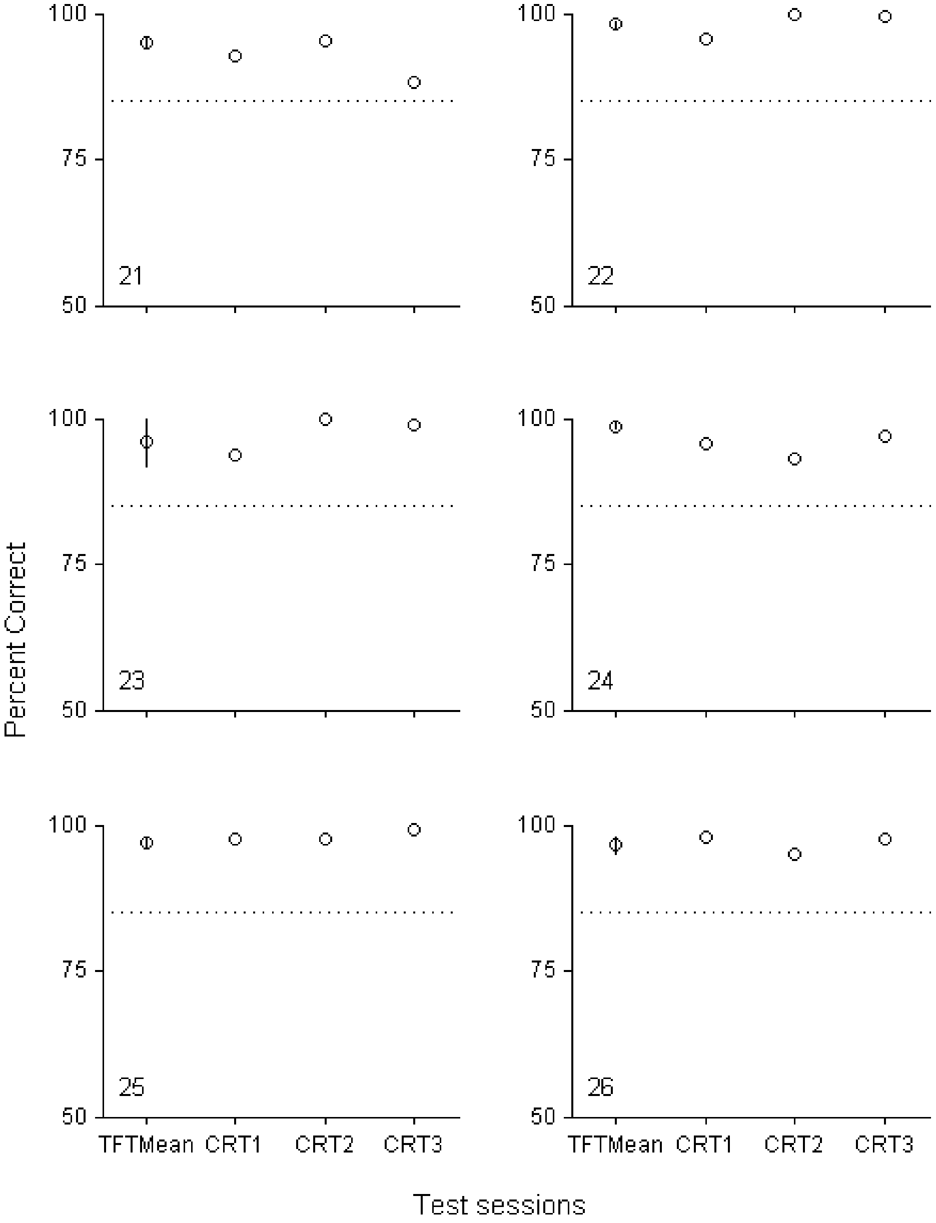


Figure 2.4. Mean percentages correct over the last five training sessions with the TFT monitor and the data from the three test sessions for all six hens (Condition 1). The vertical lines mark one standard deviation each side of the mean and the horizontal dotted lines mark 85 %.

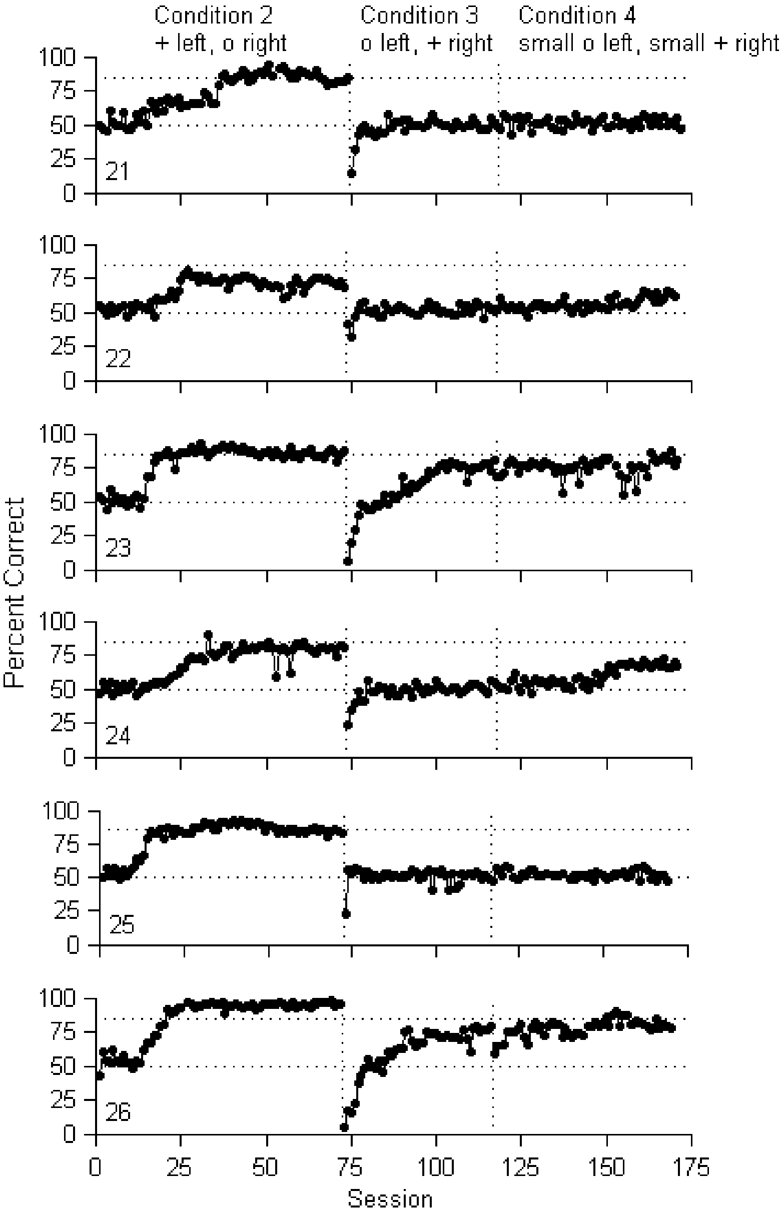


Figure 2.5a. Overall percentages correct over the training sessions for the cross and circle stimuli for Conditions 2 to 4 for all hens. The vertical dotted lines mark the condition breaks. The horizontal dotted lines mark 50 % and 85 %.

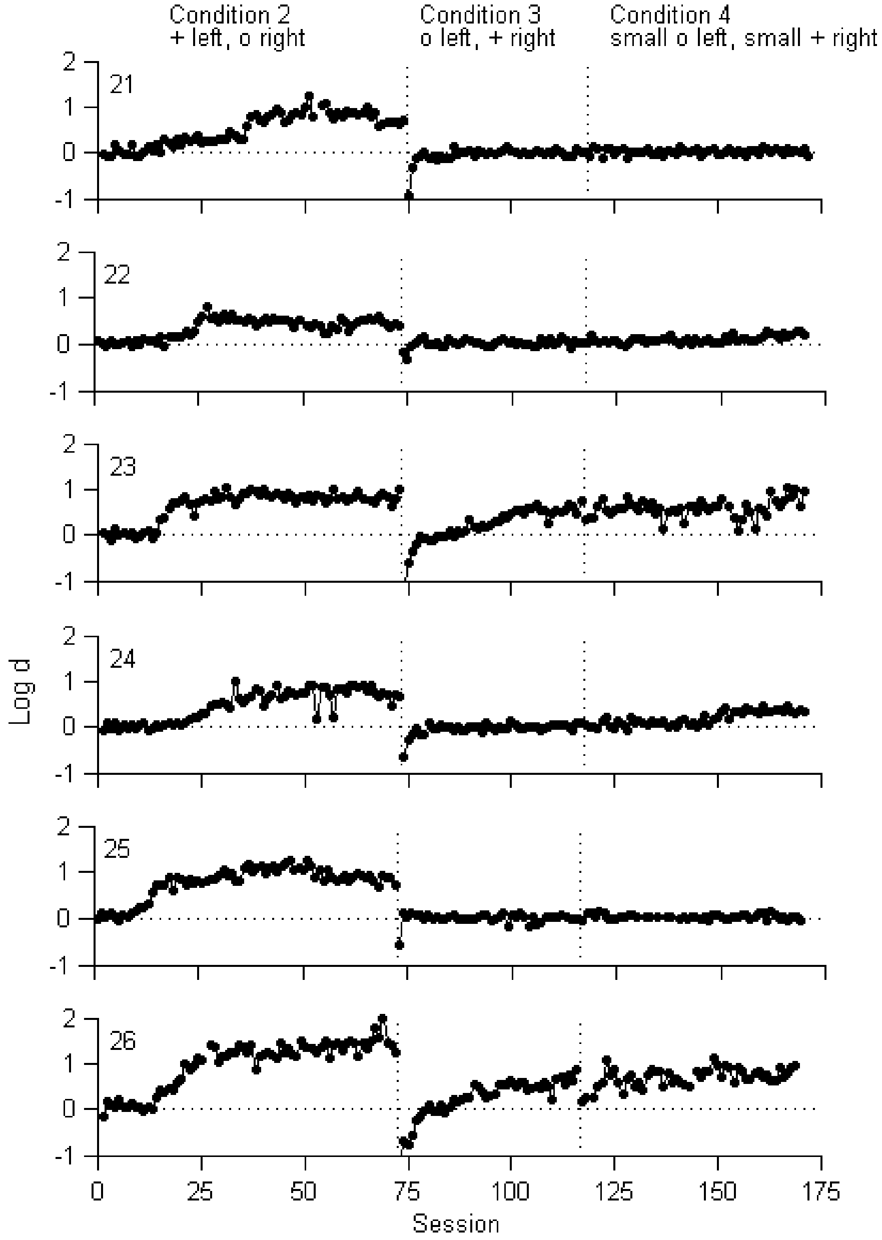


Figure 2.5b. Log *d* estimates over the training sessions with the cross and circle stimuli for Conditions 2 to 4 for all hens. The vertical dotted lines mark the condition breaks.

Condition 3. If each trial is regarded as a binomial "coin toss" then with 280 trials per session, any percentage correct over 55 % is significantly different from chance at a .05 level of significance. Hens 23 and 26's accuracies ranged from 72 to 81 % correct and 75 to 79 % correct respectively over the last five training sessions, showing their performances were significantly above chance. The remaining four hens' accuracies remained close to 50 % correct across all sessions. Condition 4 was a repeat of Condition 3, except the sizes of the stimuli were reduced. The hens received a total of 54 training sessions in this condition. Percent correct for Hens 21, 22 and 25 remained close to 50 % correct throughout this condition. Hens 23, 24 and 26's accuracies were significantly above chance (above 55 %, binomial test) in the later sessions, but was never above 85 % correct for at least five training session (range over the last five training sessions; Hen 23, 76-86 %; Hen 24, 65-72 %; Hen 26, 77-84 %).

Figures 2.6a and 2.6b present the percentage correct and log *c* estimates respectively for each stimulus plotted against session number for Conditions 2, 3 and 4 for all hens. The percent graphs show that all six hens showed response biases towards one stimulus during Condition 2. That is, all hens had a high degree of accuracy with the cross stimulus, but for Hens 22 - 25, accuracy was much lower with the circle stimulus. The log *d* estimates show that all six hens tended to peck the left key more frequently than the right, which was associated with the cross stimulus. In Condition 3, the keys associated with each stimuli were reversed, to determine if the hens' biases would also reverse (i.e., determine if the biases shown in Condition 2 were a product of selecting the key or the stimulus more often). During Condition 3, Hens 23 and 26, who had higher degree of accuracy for the cross in Condition 2, continued to do so in Condition 3 showing biased responding to the cross stimulus rather than a left key bias. This is also shown in the log *c* plots as their response biases were reversed from responding more on the left key, to responding more on the right key. The other four hens showed no consistent patterns of responding to either stimulus and no key biases. During Condition 4, Hens 23 and 26 showed a higher degree of accuracy on cross trials than on circle trials (similar to their results in Condition 3) and slightly biased responding to the right key. Hen 24 showed a higher degree of accuracy on circle trials than on cross trials, which was a reversal of her performance in Condition 2. The remaining three hens showed no differences in responding to each stimulus and no biases.

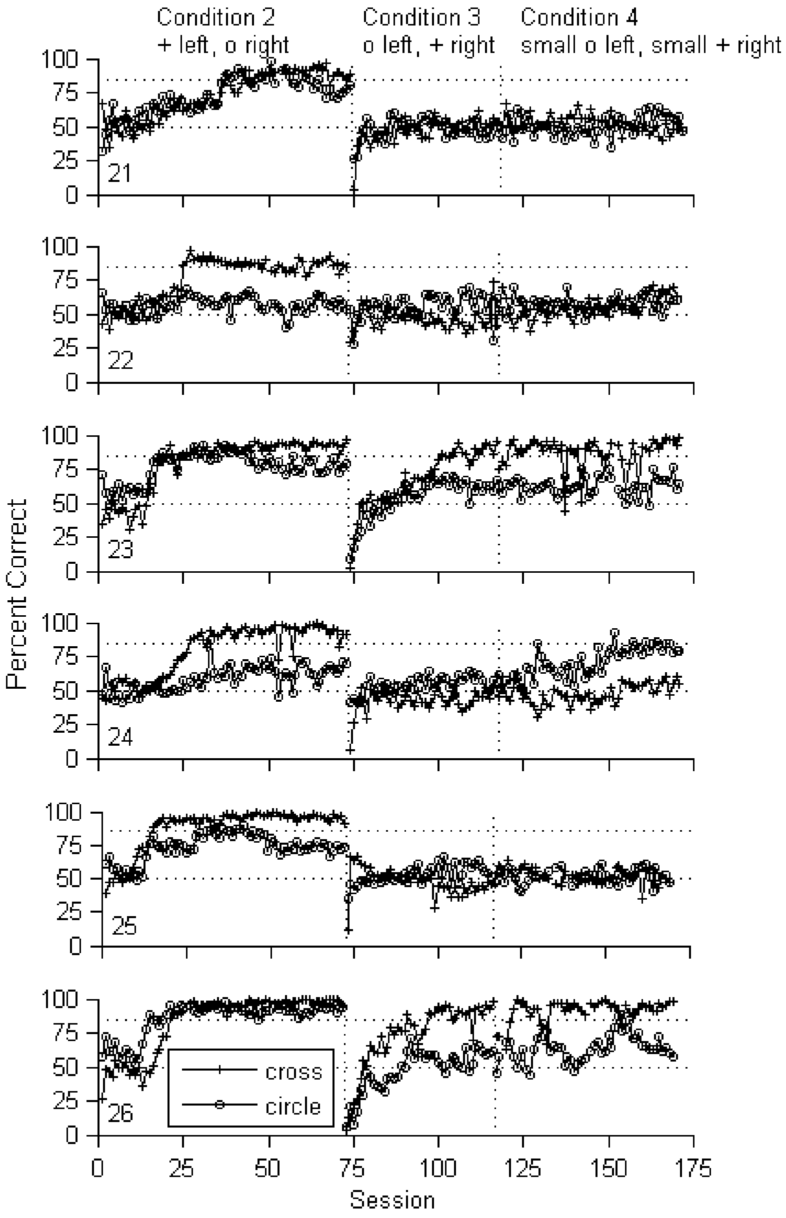


Figure 2.6a. Percentage correct for each of the stimuli over training conditions for Conditions 2 to 4 for all hens. The vertical dotted lines mark the condition breaks. The horizontal dotted lines mark 50 % and 85 %.

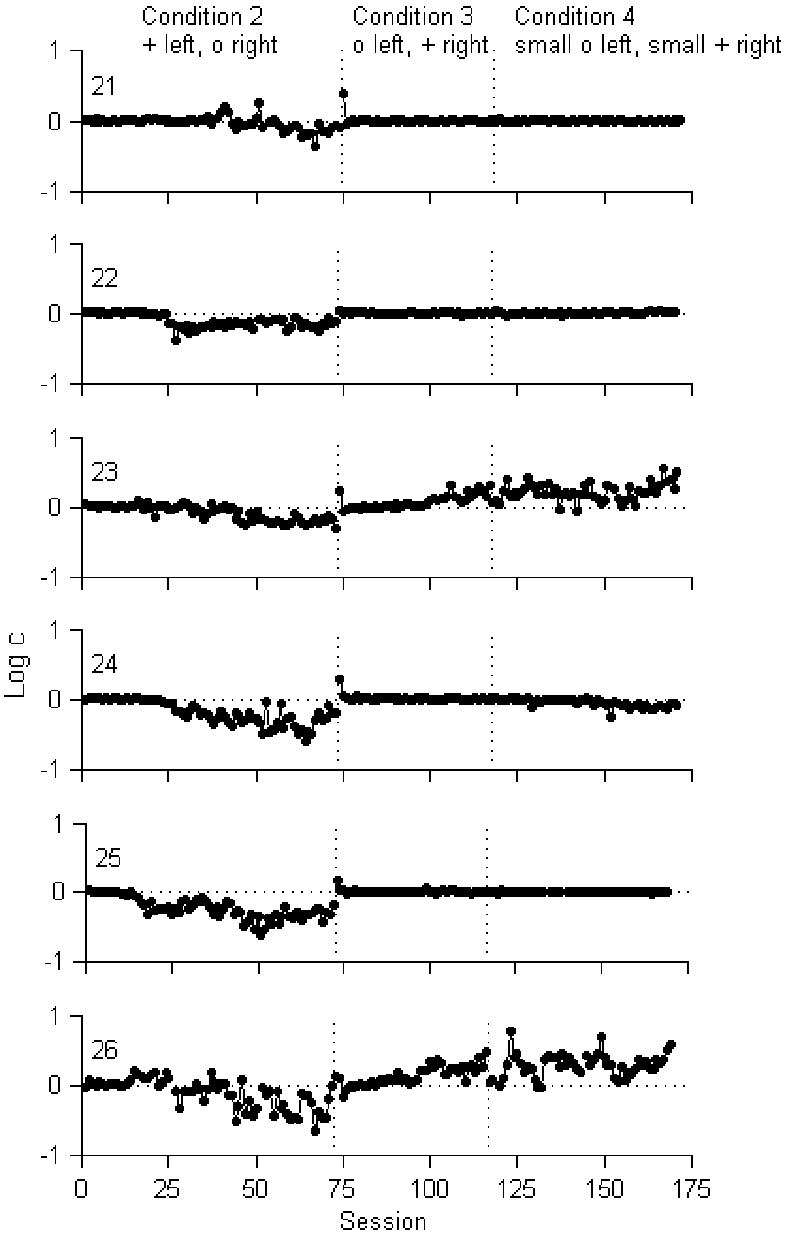


Figure 2.6b. Log *c* estimates over training conditions for Conditions 2 to 4 for all hens. The vertical dotted lines mark the condition breaks. The horizontal dotted lines mark 50 % and 85 %.

Figures 2.7a and 2.7b combines the data from Condition 4 (previously plotted in Figures 2.5a and 2.5b) and Condition 5, where the observing response was introduced. The overall percentage correct for each hen is plotted against session number in Figure 2.7a, and Log *c* estimates across session are plotted in Figure 2.7b. The locations of the test sessions are indicated by an asterisk on the x axis but data from these sessions are not plotted on these graphs. It can be seen in both graphs that the data paths for all hens' increased quickly with the introduced ratio requirement, and in all cases, accuracy increased to at least 85 % correct within 13 to 30 sessions. All hens' accuracies remained high across all sessions in Condition 5 (range, 82 to 99 % correct) with the possible exception of Hen 25 whose data show slightly more variation in accuracy than the other five hens. In addition, log *d* estimates for all hens were high in the later sessions in Condition 5.

Figures 2.8a and 2.8b present the percentage correct and log *c* estimates respectively for Conditions 4 and 5 plotted against the session number. It can be seen that, during Condition 5, there were no consistent differences in accuracy of responding across the two stimuli, and all hens generally showed high accuracy with both the cross and circle stimuli. In addition, the log *c* plots show that none of the hens showed a consistent bias for responding more to one key over the other, except for Hen 26 who showed a bias to the right key.

Figure 2.9 shows the mean percentage correct and standard deviations for the last five training sessions of Condition 5, before the test sessions with the TFT monitor. Also plotted are the data from the four test sessions for all six hens. Accuracy with both TFT test sessions with no FR requirement (test sessions 1 and 3) was lower than the training sessions with the TFT monitor. Although this difference was only significant for the second TFT test session (test session 3), they both have moderate effects (paired-samples *t*-test; training-TFT test 1, $t(5) = 2.458$, $p > .05$, $d = 1.192$; training-TFT test 3 $t(5) = 5.109$, $p < .05$, $d = 2.095$). This indicates that the change in procedure of removing the FR observing response did affect the hens' performances. Accuracy when the CRT monitor was set at 60 and 75 Hz was significantly lower (range, 56 to 84 % correct) than accuracy during the training sessions for all hens ($t(5) = 16.269$, $p < .05$, $d = 7.680$ (large effect size); $t(5) = 3.651$, $p < .05$, $d = 2.355$ (moderate effect size) respectively). It must be noted here that although the mean accuracy was lower, Hen 21 showed very high accuracy (92 % correct) with the CRT screen at 75 Hz. There was no significant difference in

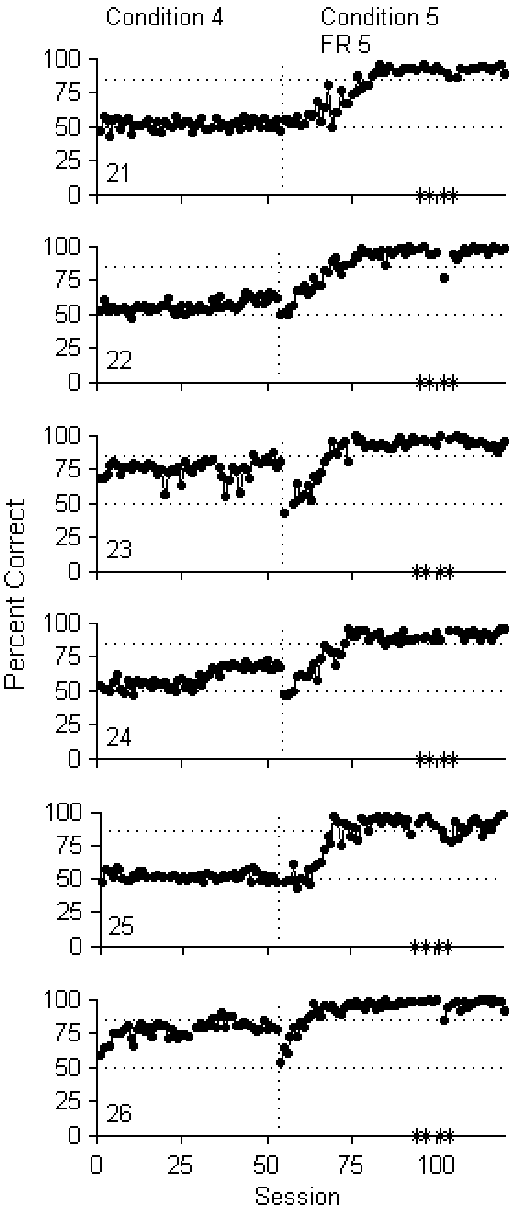


Figure 2.7a. Overall percentage correct over the training sessions for Conditions 4 and 5, for all hens. The vertical dotted line marks the condition break. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the locations of the test sessions.

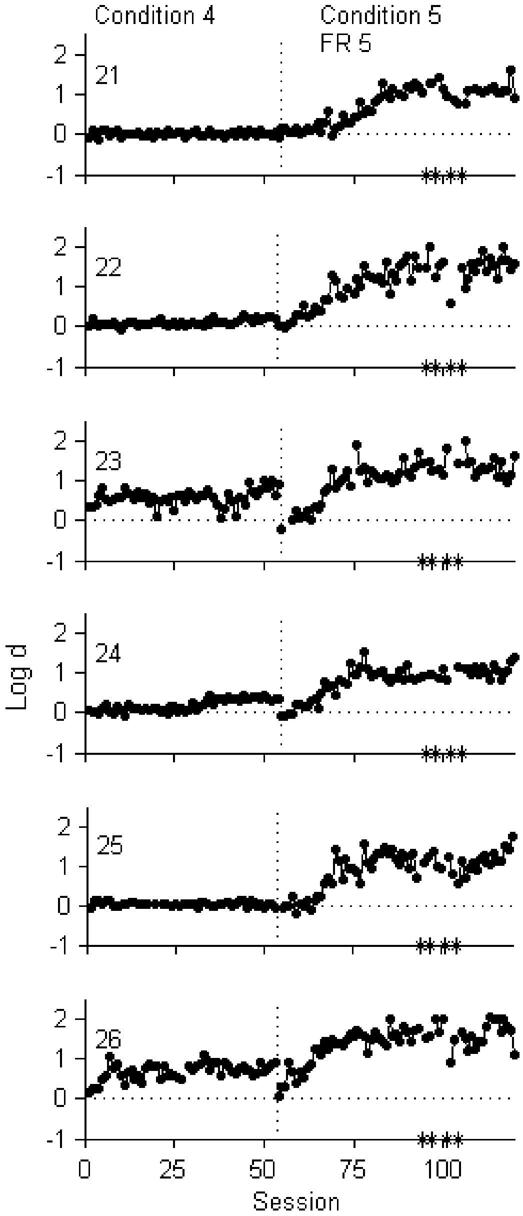


Figure 2.7b. Log *d* over the training sessions for Conditions 4 and 5, for all hens. The vertical dotted line marks the condition break. The asterisks on the x axis represent the locations of the test sessions.

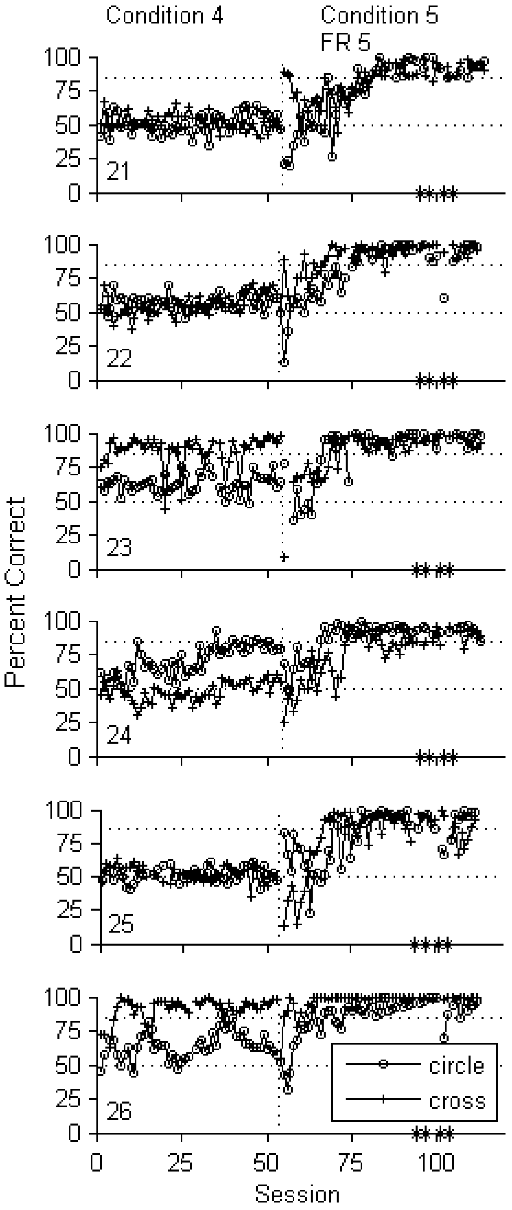


Figure 2.8a. Percentage correct for each of the stimuli over the training sessions for Conditions 4 and 5, for all hens. The vertical dotted line marks the condition break. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the locations of the test sessions.

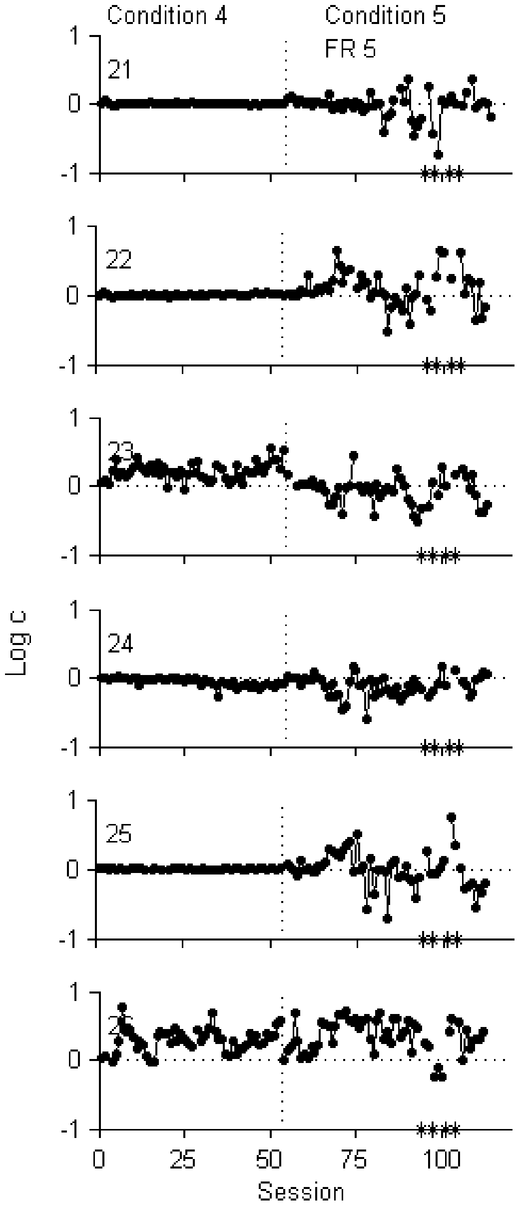


Figure 2.8b. Log *c* estimates over the training sessions for Conditions 4 and 5, for all hens. The vertical dotted line marks the condition break. The asterisks on the x axis represent the locations of the test sessions.

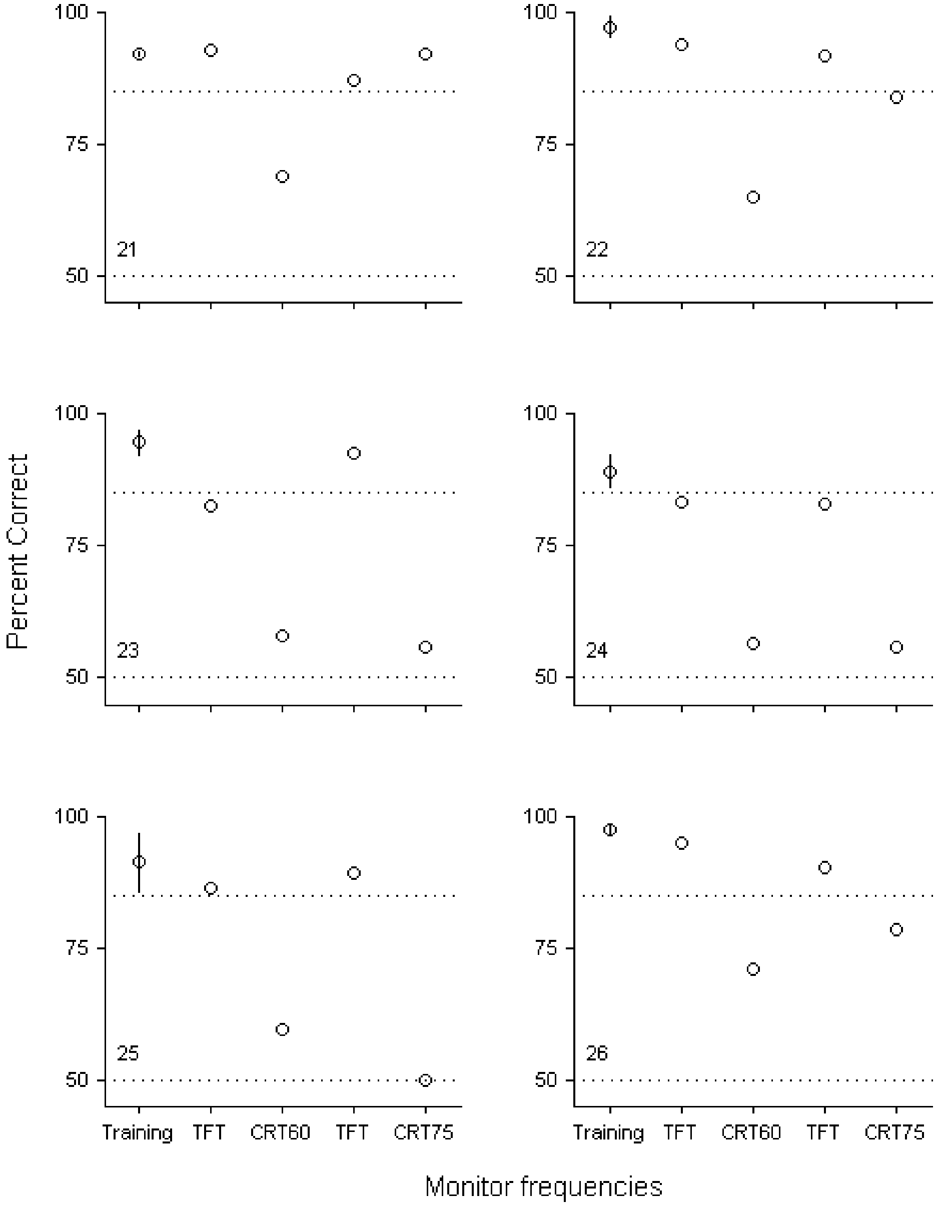


Figure 2.9. Mean percentages correct over the last five training sessions with the TFT monitor and the data from the four test sessions for all six hens (Condition 5). The vertical lines mark one standard deviation each side of the mean and the horizontal dotted lines mark 50 % and 85 %.

performance between the test sessions with the CRT screen at 65 and 70 Hz (paired-samples *t*-test, *t*(5) =-1.332, *p* >.05, *d*=-1.1) showing comparable performance across the two CRT test sessions. For each of the test sessions, percentages correct for blocks of 10 trials were examined to determine if accuracy improved during each testing session. There were 24 test sessions across all hens for this condition, and in only four of the test sessions was there any sign of an increasing trend in accuracy (Hens 21 and 25 in the first TFT test session, and Hens 21 and 23 in the 60 Hz test session). For all other test sessions, there was no visual trend in accuracy. A Theil test showed that only Hens 22 and 26's data showed a significant trend (Theil test, C = 6 for both hens, *n*= 5, *p*<.05). None of the other hens performances showed a significant trend (C = 2, 4, 5 and 4 respectively for Hens 21, 23, 24 and 25, *n*= 5, *p*>.05).

The left panel of Figures 2.10a and 2.10b shows the overall percentages correct (2.10a) and the log *d* estimates (2.10b) plotted against session number for Condition 6, for all hens. The locations of the test sessions are indicated by an asterisk on the x axis but data from these sessions are not plotted on this graph. It can be seen that all hens' accuracies and log *d* estimates began high, and remained high, likely due to previous experience with the procedure and stimuli used.

Presented on the right panel of Figures 2.10a and 2.10b are the percentages correct for each stimulus (2.10a) and log *c* estimates (2.10b) plotted against the session number for Condition 6 for all hens. It can be seen that, over the later training sessions, the hens showed no consistent differences in accuracy of responding across the two stimuli, and all hens generally showed high accuracy with both the cross and circle stimulus. None of the hens showed higher responding to one key over the other.

Figure 2.11 shows the mean percentage correct and standard deviations for the last five training sessions before test sessions with the TFT monitor. Also plotted are the data from the four test sessions (TFT and CRT set at 100 Hz). All hens were highly accurate (range, 96 to 100 % correct) with little variability during the training sessions with the TFT monitor. The third and fifth data points represent the test sessions with the TFT screen with no FR requirement. Hens 21, 24 and 26's accuracies remained high across both of these test sessions (range, 94 to 97 % correct). Hens 22, 23 and 25's accuracies dropped in the first test session (range, 75

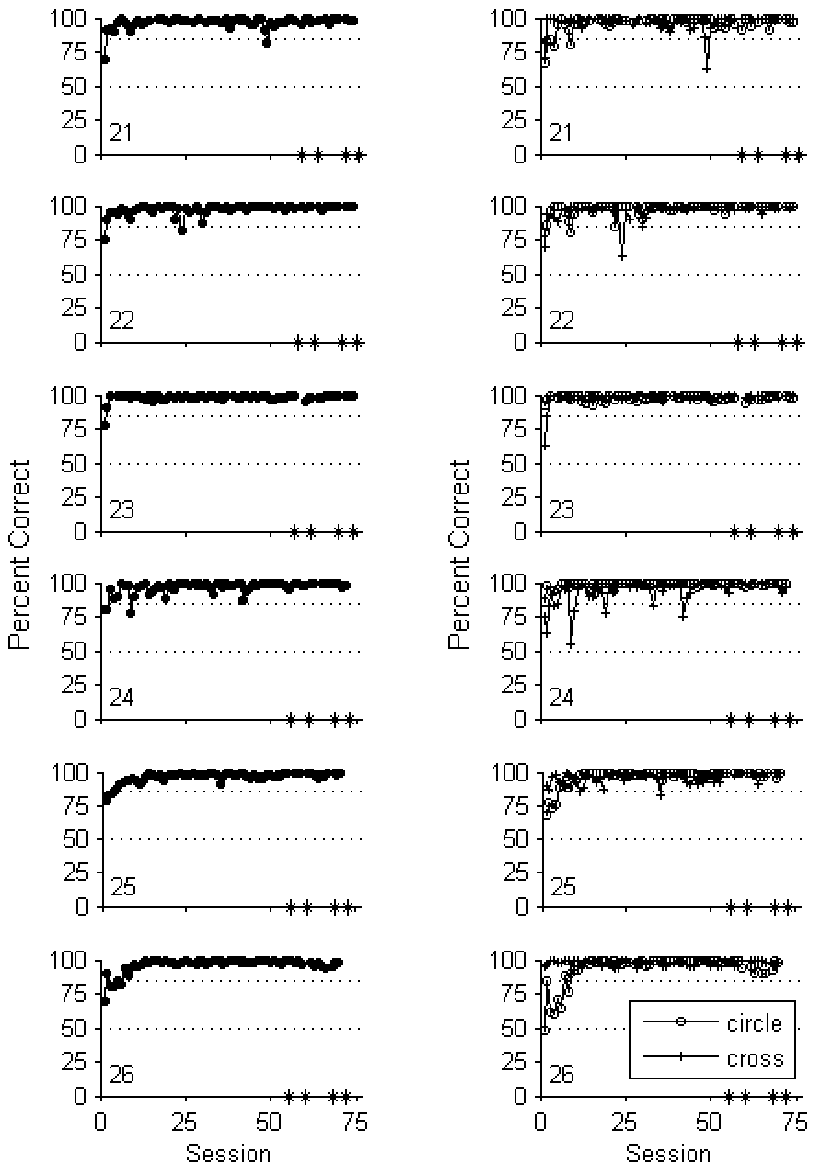


Figure 2.10a. The left panel shows the overall percentage correct over the training sessions for Condition 6, for all hens. The right panel shows the percentage correct to each stimulus. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the locations of the test sessions.

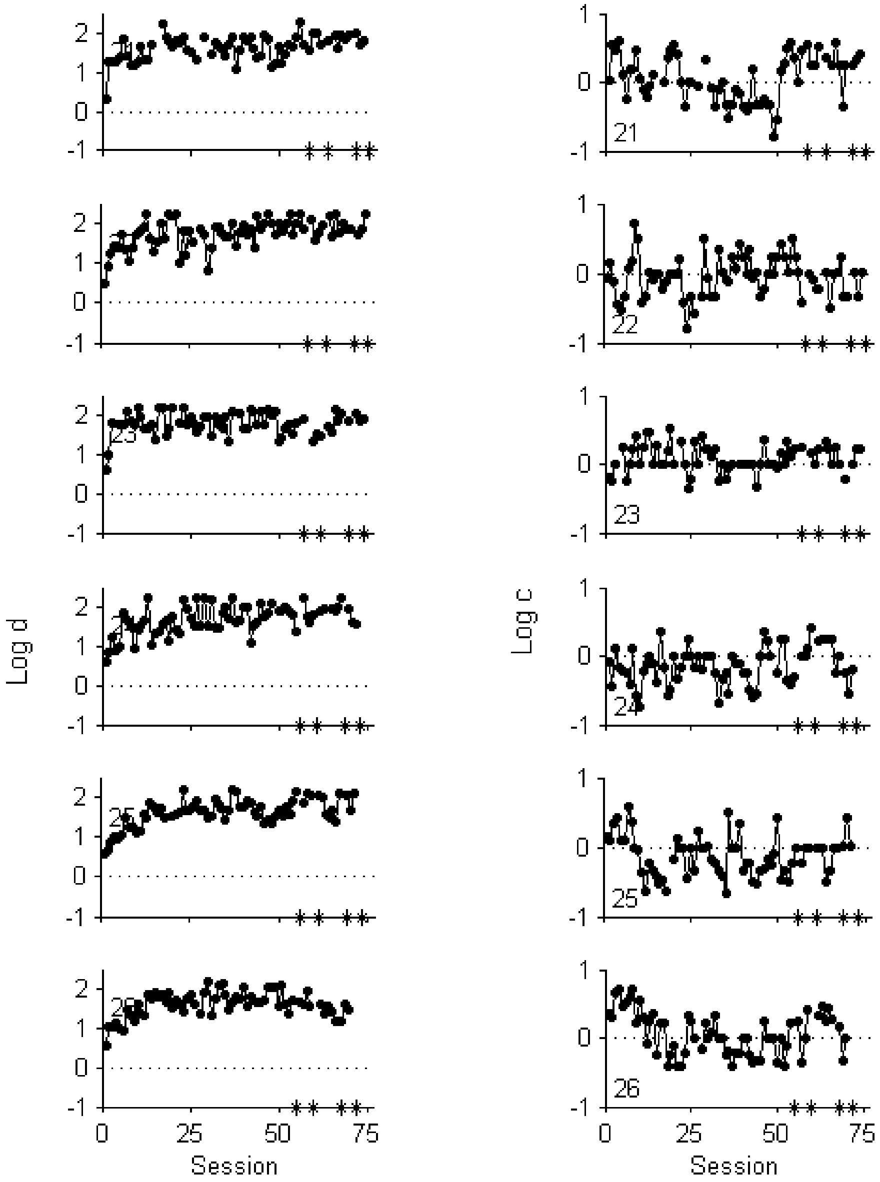


Figure 2.10b. The left panel shows the log *d* estimates over the training sessions for Condition 6, for all hens. The right panel shows the log *c* estimates over training sessions. The asterisks on the x axis represent the locations of the test sessions.

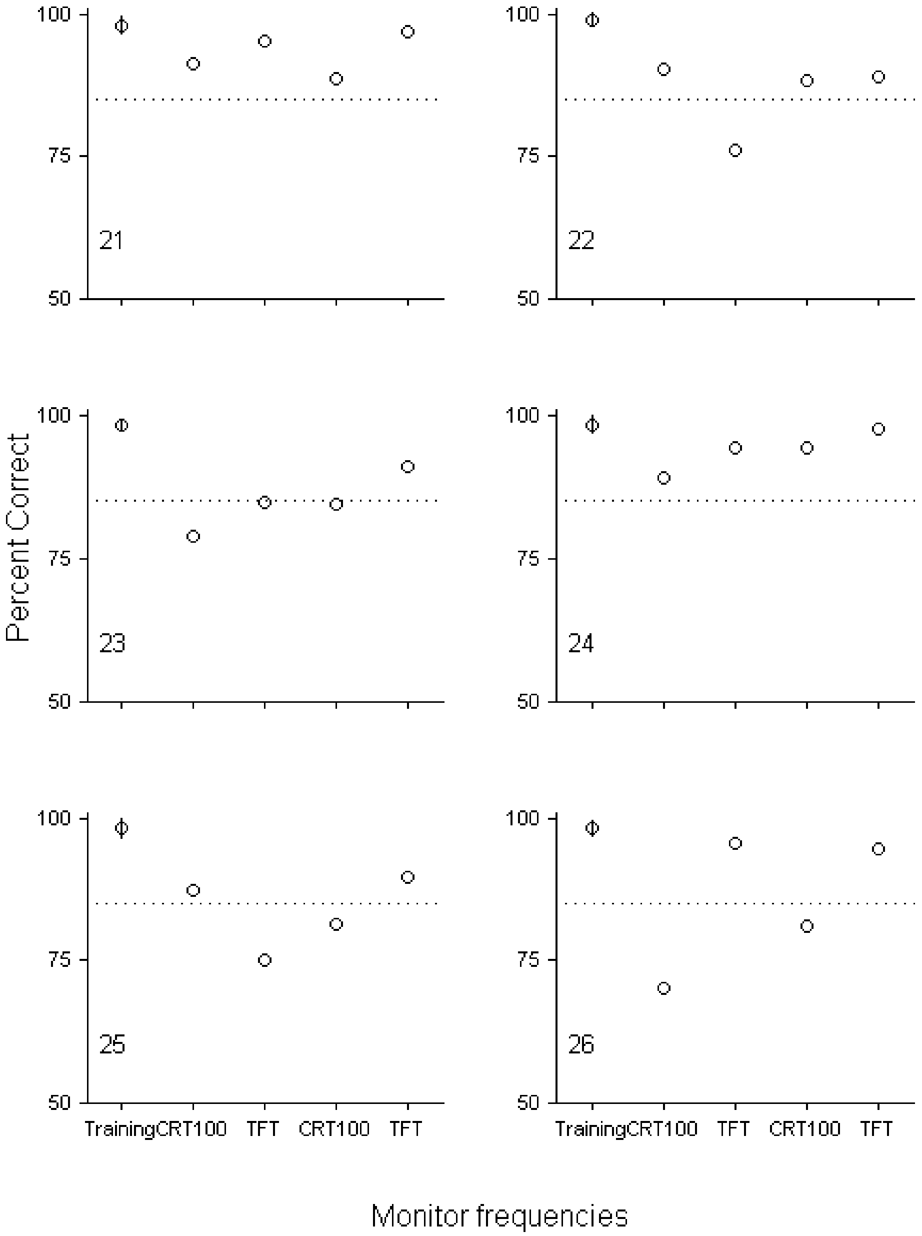


Figure 2.11. Mean percentages correct over the last five training sessions with the TFT monitor and the data from the first four test sessions for all six hens (Condition 6). The vertical lines mark one standard deviation each side of the mean and the horizontal dotted line marks 85 %.

to 85 % correct) and increased again (range, 89 to 91 % correct) in the second TFT test session. Accuracy during the second test session with the TFT screen was significantly greater than accuracy during the first test session (paired-samples *t*-test, $t(5) = -2.586$, $p < .05$, $d = -2.988$ (large effect size)). The second and fourth data points show the test sessions with the CRT screen at 100 Hz. There was no significant difference across the two 100 Hz test sessions ($t(5) = -7.19$, $p > .05$, $d = -.343$) and the average accuracy across both 100 Hz test sessions was used in further analyses. Accuracy during the CRT 100 Hz test sessions was significantly less (range, 70 to 91 % correct) than accuracy during the training sessions ($t(5) = 5.154$, $p < .05$, $d = 2.826$ (large effect size)). There were 24 test sessions across all hens for this condition, and in only three of the test sessions was there any sign of an increasing trend in accuracy (Hens 23 and 26 in the first 100 Hz and Hen 23 in the first TFT test session). For all other test sessions, there was no visual trend in accuracy. A Theil test could not be calculated as *n* was less than 4.

The mean accuracies of all test sessions at each of the monitor frequencies were averaged and are presented in Figure 2.12. It can be seen that for all hens, accuracy was highest during the training sessions. Mean accuracy across all test sessions with the TFT monitor was significantly lower (89.3 % correct) than the mean accuracy for training sessions with the TFT monitor (94.3 % correct) ($t(5) = 3.028$, $p < .05$, $d = 1.256$ (moderate effect size)). Accuracy decreased for the test sessions with the CRT monitor as the refresh rate was decreased. On average, accuracy was lowest when the CRT refresh rate was 60 Hz (Mean, 62.3) and greatest at 100 Hz (Mean, 85.3). There was no significant difference in accuracy between the 75 and 100 Hz CRT test sessions ($t(5) = -2.192$, $p > .05$, $d = 1.027$ (above RMPE according to Ferguson, 2009), but there was a significant difference in accuracy between the 60 and 100 Hz CRT test sessions ($t(5) = -5.767$, $p < .05$, $d = 3.280$ (large effect size)). There was a significant trend for Hens 22, 23, 25 and 26 (Theil test, C = 8 for all of these hens, $n = 5$, $p < .05$). There was no trend for Hens 21 and 24 (Theil test, C = 6 for these hens, $n = 5$, $p > .05$ in all cases,).

Discussion

The first aim was to assess whether hens showed similar responding with coloured stimuli shown on a TFT monitor and a CRT monitor set below their CFF. The hens learned to discriminate the coloured stimuli presented on the TFT screen,

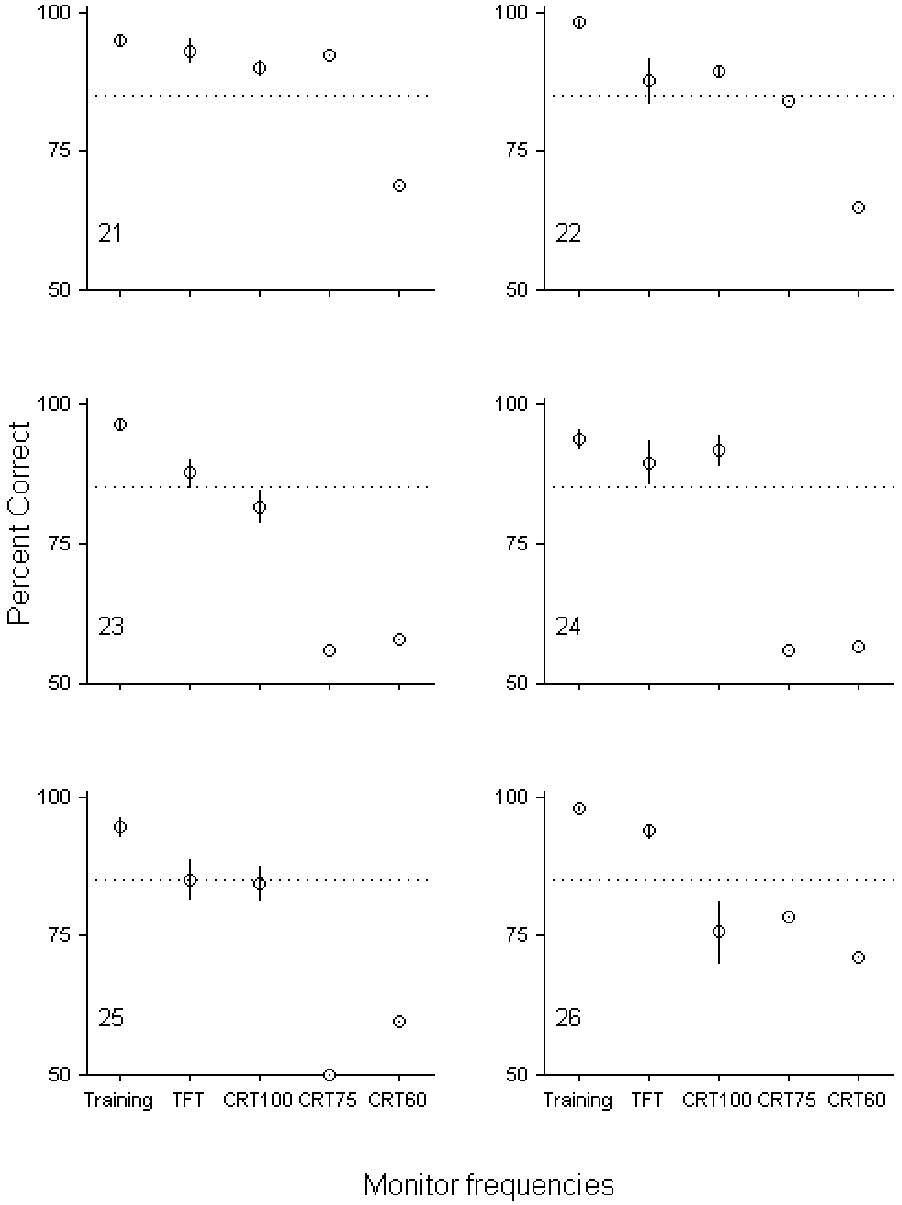


Figure 2.12. Mean percentages correct and standard error of the means for all test sessions and the data from the last five training sessions with the TFT monitor during Conditions 5 and 6, for all hens. The vertical lines marks one standard deviation each side of the mean and the horizontal dotted lines mark 85 %.

and control by the coloured stimuli transferred to the same stimuli shown on the CRT screen at 60 Hz. These results support the findings of Patterson-Kane et al. (1997) where hens transferred a discrimination between red and green cards to red and green images on a CRT screen. Both these suggest that the flicker of the CRT screen did not disrupt the hens' ability to discriminate between the coloured stimuli.

The second aim was to assess whether hens could transfer a shape discrimination learned on a TFT monitor to a CRT monitor set at different refresh rates. Once the hens had learned the discrimination, it was found that transfer with the CRT monitor was disrupted regardless of the refresh rate. In addition, accuracy decreased with the TFT test sessions suggesting that the test procedure had an effect. Nevertheless, accuracy during all of the test sessions with the CRT monitor was lower than those with the TFT monitor. These findings suggest that stimuli presented on the CRT monitor do not appear equivalent to the same stimuli presented on a TFT monitor. Overall, accuracy decreased as the refresh rate of the CRT monitor decreased (see Figure 2.12) and was lowest with the CRT monitor at 60 Hz. It is likely that the stimuli appear as flickering when presented below the hens' CFF and, therefore, may have been difficult for the hens to view.

One issue that arose during the current experiment was that, apart from three hens in Condition 2, the hens did not learn the conditional discrimination in Conditions 2 to 4 (see Figure 2.6). DeMello (1989) suggested that in conditional discrimination tasks, performance to each trial type should be examined separately to determine the pattern of responding that underlies the overall percentage correct. In conditional discrimination tasks, when reinforcement for both trial types and across both keys is kept equal (as in this experiment), equal responding to both should occur. However, responding more to one key over another shows response biases. Responding mainly on one key will result in high accuracy with one stimulus and low accuracy with the other stimulus. This can be seen in Figure 2.6 where accuracy with one stimulus is higher than accuracy on the other for some hens. During Condition 2, some of the hens developed biases towards the cross stimulus (left key) over the circle stimulus (right key). These biases in responding show the hens had learnt to behave differently to the stimuli, suggesting that they could see the stimuli as being different. Yet, they did not respond equally accurately to both stimuli. The author knows of no theoretical model of behaviour that would account for why some hens could respond accurately to one stimulus, but not to the other. If the task was simply too difficult for the hens to learn, then it seems they should not have been able to develop these biases. Yet, the hens learned the discrimination quickly when an observing response was included (FR 5), which also suggests that the task was not a difficult one for them to learn.

The development of response biases in Condition 2 (responding more to one key over the other) could result if one key was easier to peck than the other. To examine this, the key associated with each stimulus was changed. If one key was easier to peck, biases would be expected to remain high to that key regardless of which stimulus was associated with that key. For three hens, this was what happened, suggesting key biases. However, these biases were largely decreased when the keys associated with the stimuli were changed. One hen showed a change in bias to the other key, and the other two hens had no biases to start with. Taken together, the data suggest that it was not the keys that gave rise to the initial response bias.

It was considered possible that the hens' failure to learn the discrimination in Conditions 2 to 4 may have been due to the stimuli being too large. It could have been that the stimuli extended too far beyond the hens' binocular visual field to be easily discriminable. Lombardi and Delius (1990), Peissing et al. (2006) and Castro and Wasserman (2010) found that pigeons were able to generalise their discriminative performance to differently sized stimuli. Thus, the sizes of the stimuli used in the present experiment were reduced for Condition 4. It was assumed that the change in stimulus size would not affect the performance of those hens showing some discrimination (Hens 23 & 26) but may have improved the performance of the hens showing no discrimination. However, the change in stimulus size did not affect accuracy for any hens. These results suggest that the failure of the hens to discriminate the stimuli in the earlier conditions was probably not because those stimuli were too large making it difficult for the hens to view.

In Condition 5, the FR 5 requirement was included and all hens' accuracy increased quickly. The inclusion of this FR 5 meant there was an increased response requirement for each trial and that the hens were in the presence of the stimulus for longer. When DeMello, Foster and Temple (1993) changed the response requirements from FR1 to FR5 schedules, they found that hens' accuracy on a visual acuity task increased. In addition, White (1985) found that increasing an FR requirement from 1 to 5 increased pigeons' accuracy in a delayed-matching-to-sample (DMTS) task. Both studies show that increasing the response requirement can result in increased accuracy as was found here.

This finding of increased accuracy shows that the task was not too difficult for the hens to learn. The location of the observing response was directly below the discriminative stimuli and it may have served to orient the hens towards the stimuli, leading to higher accuracy. Alternatively, Sacks, Kamil and Mack (1972) stated that increasing the response requirement increases the amount of effort required on each trial, ensuring that any incorrect responses were now more costly, which would lead to increased accuracy. Also, increasing the response requirement also increased the duration of the stimulus presentation and this alone might have lead to increased accuracy. Any, or all of these factors, may have contributed to the increased accuracy found in this study.

There is a study that attempted to determine whether it was sample duration or response requirement that affected accuracy in a delayed-matching-to-sample discrimination. Foster, Temple, Mackenzie, DeMello and Poling (1995) controlled both the duration of the sample stimulus and the size of the response requirement. They found that both sample duration and response requirement directly and independently affected accuracy. Their results suggest that it may be either the increased response requirement or duration of stimulus presentation (or possibly both) that led to increased accuracy in the present experiment. However, it must be noted that they presented the stimuli on the response keys, while in the present experiment the response requirement was on the same screen, but in a different location to the stimuli. It seems reasonable to assume that the procedures were similar enough that similar processes may be involved.

This experiment established a possible procedure to test for correspondence across different stimulus presentation methods. The hens were able to learn a conditional discrimination task which included an FR 5 observing response. The data suggests that the decrease in accuracy with the CRT monitors was a result of the increase in flicker as the refresh rates were changed. However, there was a confound with the present experiment as there was no FR requirement during the test sessions as the CRT monitor could not record pecks. Accuracy was lower for the test sessions with the TFT test sessions in Condition 5 where there was no FR requirement, implying that the change in procedure also affected the hens' discrimination. It would obviously be useful to access technology that could record pecks to both the

TFT and CRT monitors, allowing the training and transfer sessions to be identical. This would ensure that any changes in responding were due to the change in monitor and refresh rate, rather than the change in testing procedure.

EXPERIMENT 3 [2]

The previous experiment showed that hens could transfer a learned discrimination between coloured stimuli presented on a TFT monitor to a CRT monitor set at 60 Hz. The degree of transfer of a learned discrimination between two shapes (cross and circle) presented on a TFT monitor to a CRT monitor depended on the refresh rate. In general, the transfer was high at the highest refresh rate (100 Hz), and was low at the lowest refresh rate (60 Hz).

One problem with interpreting the results of the previous experiment was that a procedural change for the last condition (i.e., the inclusion of an observing response on the TFT monitor) meant that the testing procedure had to be different from the training procedure. Before the start of this next experiment, infrared screens were found that allowed accurate detection of the location of a peck on a surface without the need for a touch screen. These infrared screens could be attached around the front of the CRT and TFT monitors, and could detect the number and location of pecks to a screen, thereby eliminating the need to alter the procedure during test sessions.

The next experiment aimed to replicate and extend the previous study using the infrared screens to keep testing and training similar and to help make sure the observing response occurred. Thus, the present study used this technology to investigate how altering the refresh rate of a CRT monitor over a series of test sessions, over a range from above to below hens' CFF, affected their discriminative performance.

As reinforcement was still available during test sessions, the hens received only one test session at each of the CRT refresh rates (except for one additional 100 Hz session). This was done as increases in accuracy across many test sessions would probably result from the hens learning a new discrimination, rather than transferring the previously learned discrimination.

Several different conditional discriminations were used. In Condition 1 the stimuli were the black cross and circle from Experiment 2. This allowed the methods used in the two experiments to be compared. In Conditions 2 and 3 Lego™ shapes were used as stimuli. The selection of these came from experience with the same shapes used in an experiment which had already been completed and is presented

---

[2] Parts of this experiment have been published as Railton, R.C.R., Foster, T.M., & Temple, W. (2010). Transfer of stimulus control from a TFT to CRT screen. *Behavioural Processes, 85,* 111-115.

later in this thesis as Experiment 7. The results from that experiment showed these stimuli to be a difficult, but possible, discrimination for hens. However, after failure to gain discrimination using these Lego^{TM} shapes, Condition 4 involved a colour discrimination. As these hens had had previous experience discriminating between green and red stimuli, blue and yellow were used in this experiment. As stated previously, hens have been shown to discriminate easily between colours including blue and yellow. For example, Huber-Eicher (2004) and Zupan, Kruschwitz and Huber-Eicher (2007) found hens laid more eggs in yellow coloured nest boxes, over blue, green or red nest boxes. In addition, Jones and Carmichael (1998) found that hens pecked more often at white or yellow, rather than blue or orange, bunches of string, showing that the hens could distinguish between these colours.

In Condition 5, the same Lego^{TM} shapes used in Conditions 2 and 3 were used, however, an attempt was made to make them easier to discriminate. This involved removal of the 3D cues (i.e., sides of the shapes) from the images. The stimuli used in Condition 6 were black-on-white line drawings of a watering can and an iron. These stimuli were selected as pigeons had been shown to successfully discriminate between them in a study by Wasserman, Kirkpatrick-Steger, Van Hamme and Biederman (1993) assessing spatial organisation.

Method

Subjects

The subjects were five of the hens used in Experiment 2. Hen 24 died of lymphoid leucosis after Experiment 2 and was not replaced.

Apparatus

The apparatus was the same as that used in Experiment 2 except that infrared screens (IR Touchscreen 12 in. USB) were attached to the front of the TFT and CRT monitors. The infrared screens were large enough that their edges lay outside the opening in the chamber wall. The infrared screens were controlled by the main experimental programme and a purpose built interface which linked into the MED programme, and the location and number of pecks could be recorded.

Procedure

Discrimination Training. The general procedure was the same as used in Condition 5 of Experiment 2. Table 3.1 presents a list of the conditions, the stimuli

Table 3.1.

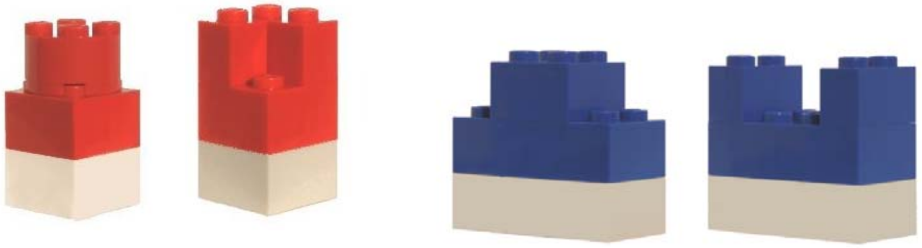
The order of experimental conditions, together with the stimuli used, number of sessions over which the VR was increased, the final VR, sessions to VR, the number of sessions in each condition, and the number of hens responding at or above 85 % for at least five sessions.

Condition	Stimuli associated with left key	Stimuli associated with right key	Final VR	Sessions to VR	No. of sessions	No. of hens at or above 85%
1	Circle	Cross	5	0	95	5
2	Shape 1	Shape 2	4	14	31	0
3	U-shape	Inverted T-shape	4	40	47	0
4	Blue square	Yellow square	5	15	95	5
5	2D U-shape	2D Inverted T-shape	5	32	59	0
6	Iron	Watering Can	5	23	69	4

used, number of sessions over which the VR was increased, sessions to VR, the final VR requirement, the number of sessions in each condition, and the number of hens responding at or above 85 % for at least five sessions. Decisions to increase the VR were based on a combination of relatively stable performance and the subject having received approximately 30 reinforcers during the experimental session. The final VR used was either VR 4 or VR 5.

Condition 1 followed immediately from the last condition in Experiment 2, and so a VR 5 was already in effect. The discriminative stimuli used in Condition 1 were the same as those used in Condition 5 of Experiment 2 (the cross and circle). The discriminative stimuli used in Conditions 2 to 6 are presented in Figure 3.1. In Condition 2, discriminative stimuli were pictures of red and white Lego^{TM} blocks (21-mm wide × 32-mm high). Each stimulus had a white square base with a red square clipped onto it. On top of this was either a red circular block (termed shape 1) or a cubed block with one corner missing (termed shape 2). The stimuli used in Condition 3 were pictures of blue and white Lego^{TM} blocks (31-mm wide × 31-mm high). Each stimulus consisted of a picture of a blue rectangular block clipped on top of a white rectangular block. On top of this was either a square blue block clipped in the middle of the rectangular blocks (termed the inverted T-shape), or two small blue rectangular blocks clipped to each end of the rectangular blocks (termed the U-shape). During Condition 4, the stimuli were a blue and a yellow square (45-mm wide × 47-mm high) shown against a white background. The stimuli used in Condition 5 were the same pictures of blue and white Lego^{TM} blocks used in Condition 3 (i.e., the U-shape and inverted T-shape), with the sides and shading deleted to remove any three-dimensional cues. During Condition 6, the stimuli used were line drawings of an iron (63-mm wide × 28-mm high) or a watering can (46-mm wide × 44-mm high). For all conditions, the stimuli were presented against a white background.

Sessions were terminated after 40 min had elapsed or after 30 reinforcers had been obtained, whichever occurred first. Test sessions occurred when at least two hens' accuracy was at or above 85 % for at least five, not necessarily consecutive, sessions. None of the hens' performances reached 85 % correct in Conditions 2, 3 or 5; therefore test sessions occurred only in Conditions 1, 4, and 6. During Condition 6, Hen 25 was not responding above 50 % before test sessions began, and her test session data are excluded from the analyses for this condition.

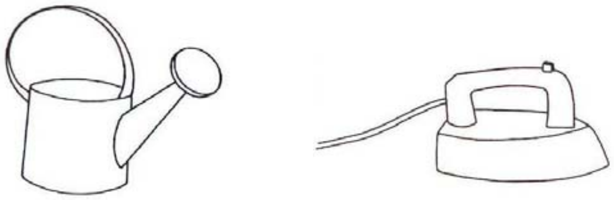


Shape 1 Shape 2 Inverted T-shape U-shape

Condtion 2 Condition 3

Yellow square Blue square Inverted 2D T-shape 2D U-shape

Condition 4 Condition 5



Watering Can Iron

Condition 6

Figure 3.1. Images of the stimuli used in Conditions 2-6.

Test sessions. The test sessions followed the same procedure as the discrimination training, except that the stimuli were presented on a CRT monitor rather than a TFT monitor. There were five test sessions across each condition. The refresh rate of the CRT monitor was set at 100, 85, 75, and 60 Hz over the first four test sessions and then at 100 Hz for the fifth test session. After each test session, a hen received at least two training sessions with the TFT monitor. Hen 26 responded atypically in the first test session at 75 Hz in Condition 1, and was given additional test sessions at 75 and 85 Hz, but was not exposed to the CRT monitor at 60 Hz.

Results

Sometimes data points are missing as data were discarded if an egg was laid during an experimental session, except for Hen 23, when there were days when she did not respond.

Figures 3.2a and 3.2b shows the overall percentage correct with the shape stimuli (cross and circle) and log *d* estimates plotted against session number for all hens (Condition 1). The locations of the test sessions are indicated by the asterisks on the x axis, and data from these sessions are not plotted. Due to equipment problems, there were a large number of training sessions between the first four and last test sessions. Hen 23 had health problems when the other hens completed test session number 4, and so her fourth test session was conducted after that of the other hens. Condition 1 followed immediately from Experiment 2 and it can be seen in both graphs that all hens were highly accurate from the start (over 95 % correct) and their performances remained stable throughout the training sessions. Performance was not disrupted by the test sessions. Test sessions could be started immediately as all hens were responding above 85 % correct at the start of this condition. Figure 3.3a and 3.3b presents the percentages correct to each stimulus and log *c* estimates plotted against session number. These graphs shows there were no consistent differences in accuracy across the two stimuli, except for Hen 26 who showed more accurate responding on the right key.

Figure 3.4 shows the mean percentages correct and standard deviations over the last five training sessions with the TFT monitor with the cross and circle. As the test sessions for this condition began immediately after the last condition of Experiment 2, the training session data are the same as those plotted in Figure 2.11, and are close to 100 % correct with very little variability for all hens. Also shown in

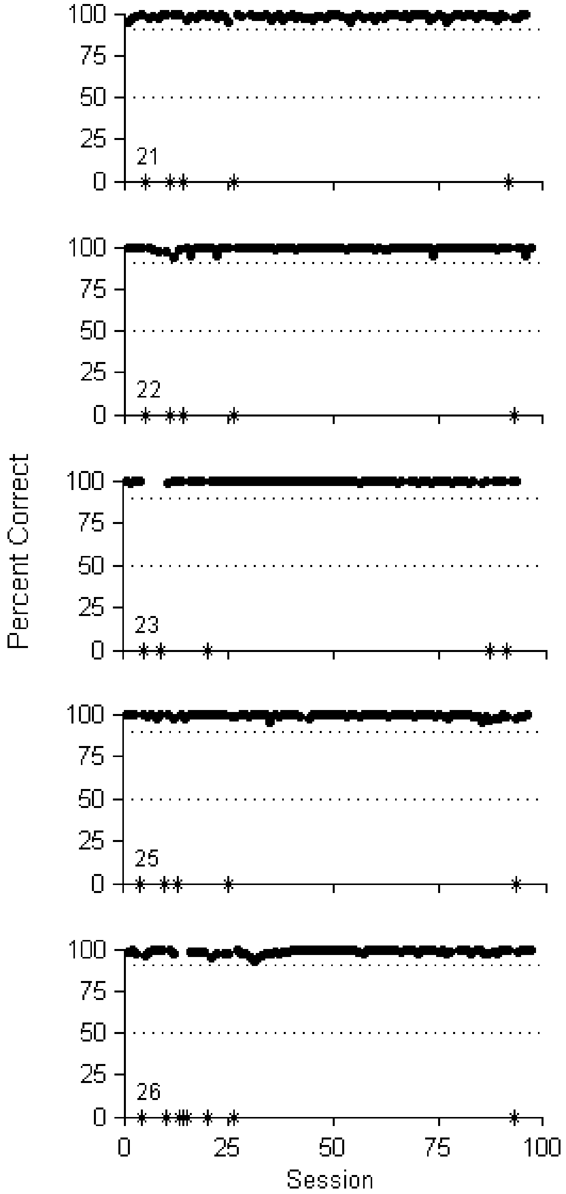


Figure 3.2a. Overall percentage correct to the shape stimuli (cross and circle) for all five hens plotted against session number for all hens (Condition 1). The horizontal lines mark 50 % and 85 %. The asterisks on the x axis represent the locations of the test sessions.

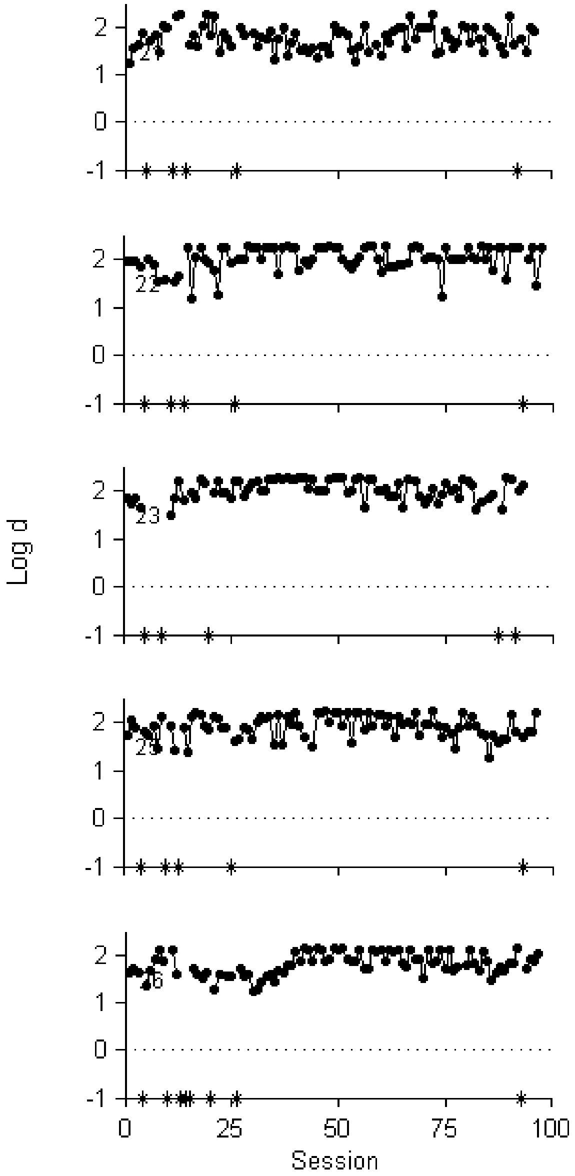


Figure 3.2b. Log *d* estimates for all five hens plotted against session number for all hens (Condition 1). The asterisks on the x axis represent the locations of the test sessions.

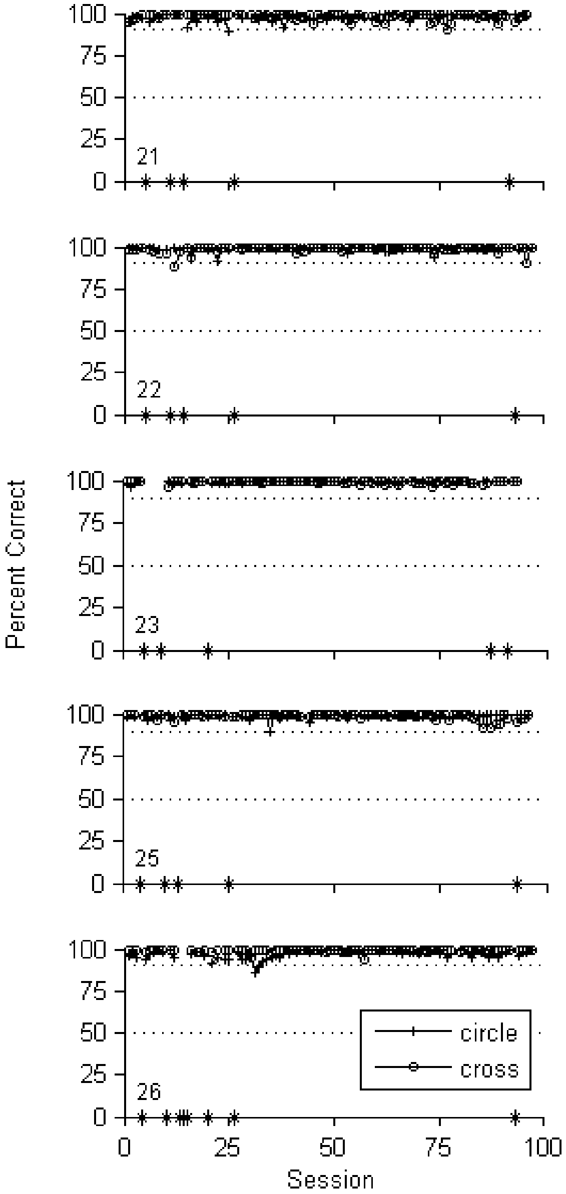


Figure 3.3a. Percentage correct for each of the shape stimuli over the training sessions plotted against session number for all hens (Condition 1). The horizontal dotted lines mark 50 and 85 %. The asterisks on the x axis represent the locations of the test sessions.

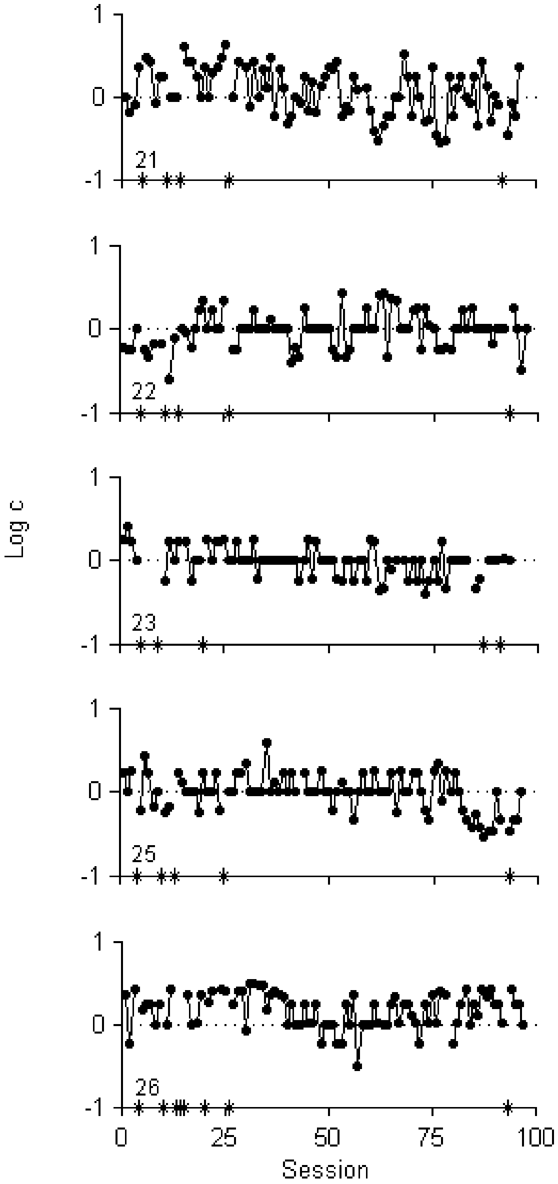


Figure 3.3b. Log *c* estimates plotted against session number for all hens (Condition 1). The asterisks on the x axis represent the locations of the test sessions.

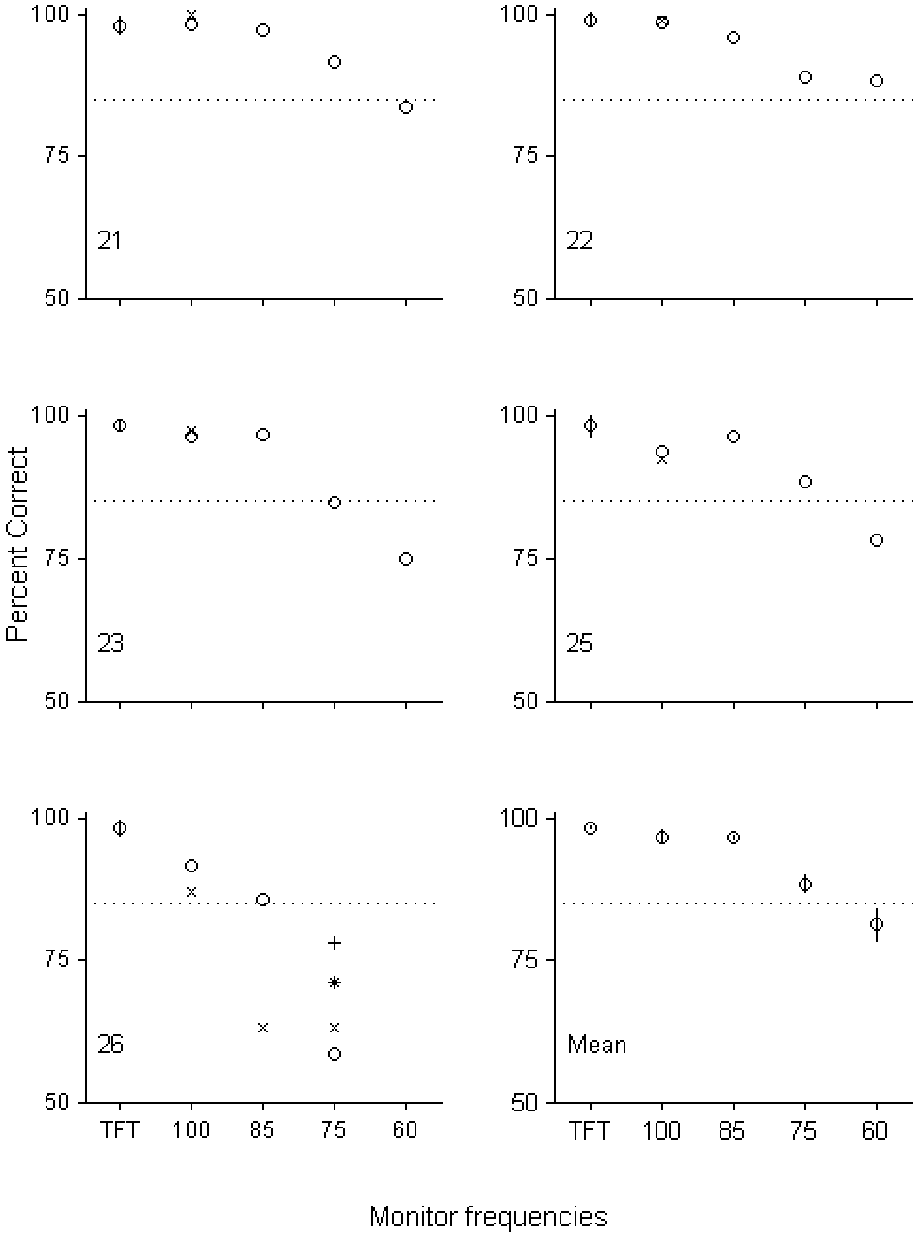


Figure 3.4. Mean percentages correct and standard deviation over the last five training sessions (TFT) and the data from the test sessions for the shape stimuli for all hens (Condition 1). The vertical lines mark one standard deviation each side of the mean and the dotted horizontal line shows 85 %. The replication of the 100 Hz condition is shown as an x. The additional symbols are explained in the text. The bottom right panel shows the average percentage correct and standard error of the means at each frequency level across all five hens.

Figure 3.4 are the data from the five test sessions. The data from the last (100 Hz) test session is shown as a cross. The bottom right panel shows the mean percentage correct and standard error of the means at each frequency level across all five hens. The data from the extra test sessions completed by Hen 26 are plotted, but were not included in the mean percentage correct or the statistical analyses outlined below. For this hen, the cross represents the data for the second test session at each monitor frequency, the plus sign represents the data from the third test session and the asterisk represents the data from the fourth test session at 75 Hz. On average, accuracy decreased across hens as the refresh rate of the CRT monitor was decreased.

Statistical analysis showed no significant difference between the data from the first and last100 Hz test sessions (paired-samples *t*-test, $t(3) = -0.840$, $p > .05$, d=-.639). Hence, the averages of the 100 Hz test conditions were used for all further analyses. Accuracy decreased as the refresh rate of the CRT monitor was decreased.

A repeated-measures ANOVA over the data from the test sessions showed the change was statistically significant ($F(4,12) = 29.08$, $p<.05$, $\eta^2=.906$ (strong effect size)). The data for all hens showed a significant trend (Theil test, C = 8, 7, 10, 8 and 6, for Hens 21 to 26 respectively, n= 5, $p>.05$ in all cases,). Percentages correct at 100 Hz were only slightly lower than those from the TFT training sessions and were not significantly different from the TFT data ($t(3) = 1.515$, $p > .05$, d=1.111). However, percentage correct when the CRT monitor was set at 85, 75 and 60 Hz was significantly lower than that when the TFT monitor was used in training ($t(3) = 3.544$, $p < .05$, d=1.731 (moderate effect size); $t(3) = 6.340$, $p < .05$, d=3.417 (strong effect size); $t(3) = 6.245$, $p < .05$, d=4.641 (strong effect size) respectively). There was no significant difference between data when the CRT monitor was set at 100 Hz and 85 Hz ($t(3) = .290$, $p > .05$, d=.15). However, percentage correct with the CRT monitor set at 100 Hz was significantly different from that at 75 and 60 Hz ($t(3) = 5.455$, $p < .05$, d=2.723 (strong effect size); $t(3) = 6.716$, $p < .05$, d=4.130 (strong effect size) respectively). In addition, there was a significant difference between percentage correct with the CRT monitor set at 85 and 75 Hz, 85 and 60 Hz, and 75 and 60 Hz ($t(3) = 6.203$, $p < .05$, d=4.950 (strong effect size); $t(3) = 5.127$, $p < .05$, d=3.148 (strong effect size); $t(3) = 3.258$, $p < .05$, d=2.079 (moderate effect size) respectively).

For each of the test sessions, percentages correct for blocks of 10 trials were examined to determine if accuracy changed during each testing session. There were

25 test sessions across all hens for this condition, and in only four of the test sessions was there any sign of an increasing trend in accuracy (Hens 22 and 23 in the 75 Hz test session, and Hens 21 and 22 in the 60 Hz test session). For all other test sessions, there was no visual trend in accuracy.

On the left panel of Figures 3.5a and 3.5b are the overall percentage correct with the red and white Lego^{TM} blocks (Condition 2) and the log *d* estimates, plotted against session number. Both graphs show that the hens' accuracy remained around 50 % (i.e., at chance levels) after 29 to 31 training sessions. No hens' performances were trending or reached 85 % correct. The right panel of Figures 3.5a and 3.5b presents the percentage correct to each stimulus and the log *c* estimates over sessions. Hen 21 and 22's responding were around 50 % correct with both of the stimuli. None of the hens showed biased responding to one key over the other.

The left panel of Figures 3.6a and 3.6b presents the overall percentage correct for all sessions with the blue and white Lego^{TM} blocks (Condition 3), and the log *d* estimates plotted against session number. The data show that all hens' performances remained near 50 % correct over all training sessions and discriminative performance was low. The right panel of Figures 3.6a and 3.6b presents the percentage correct for each stimulus in Condition 3 and the log *c* estimates over sessions. None of the hens show consistently higher accuracy with either of the stimuli over the later training sessions, and show no biases to selecting either key.

Figures 3.7a and 3.7b show the overall percentage correct over all training sessions for the colour discrimination (Condition 4), and the log *d* estimates plotted against session number. The locations of the test sessions are indicated by asterisks on the x axis. All hens' accuracies were at or above 85 % within 28 training sessions (shown by the vertical dotted line). For all hens (except Hen 23), performances remained stable and at or above 85 % for all remaining training sessions, including those before and after each test session. Hen 23's performance was generally stable and above 85 % correct but after the final test session her performance became more variable. Between every test session, each hen had a minimum of three training sessions with the TFT monitor. Accuracy for these interspersed sessions ranged from 89 to 100 %. The data path for the log *d* estimates show all hens generally show an increasing trend.

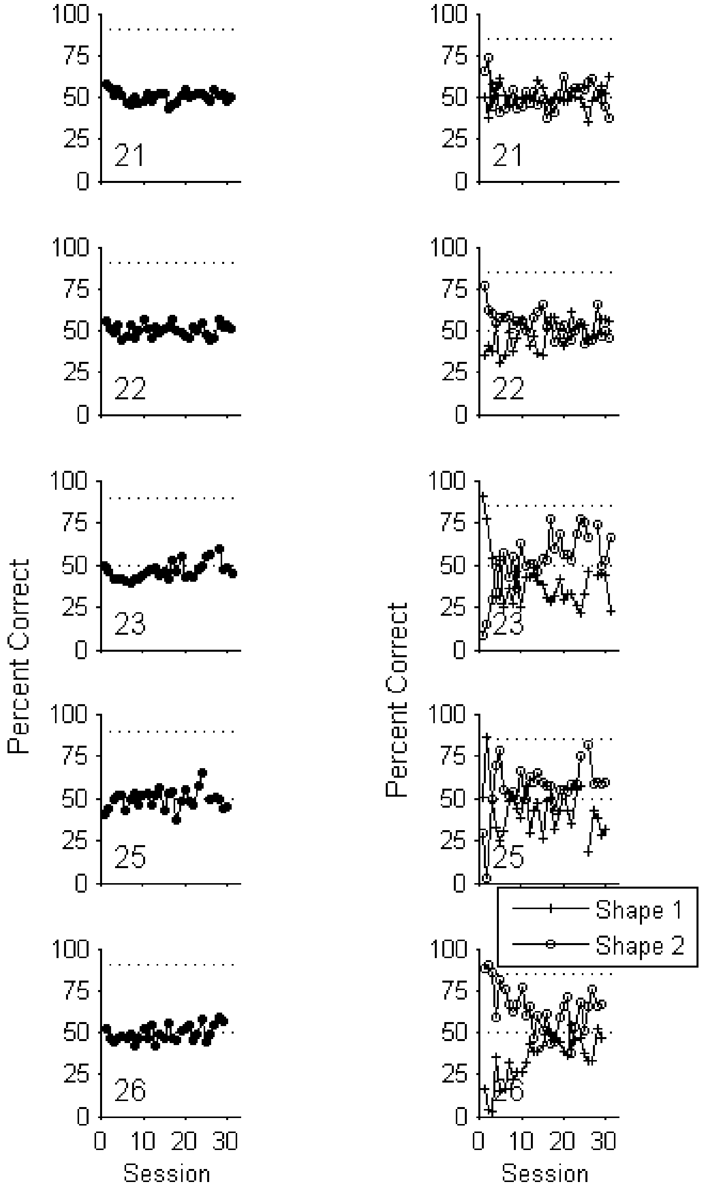


Figure 3.5a. The left panel shows the overall percentage correct with the red and white Lego™ blocks (Condition 2) for all five hens plotted against session number. The right panel shows the percentage correct to each of the red and white Lego™ stimuli. The horizontal dotted lines mark 50 and 85 %.

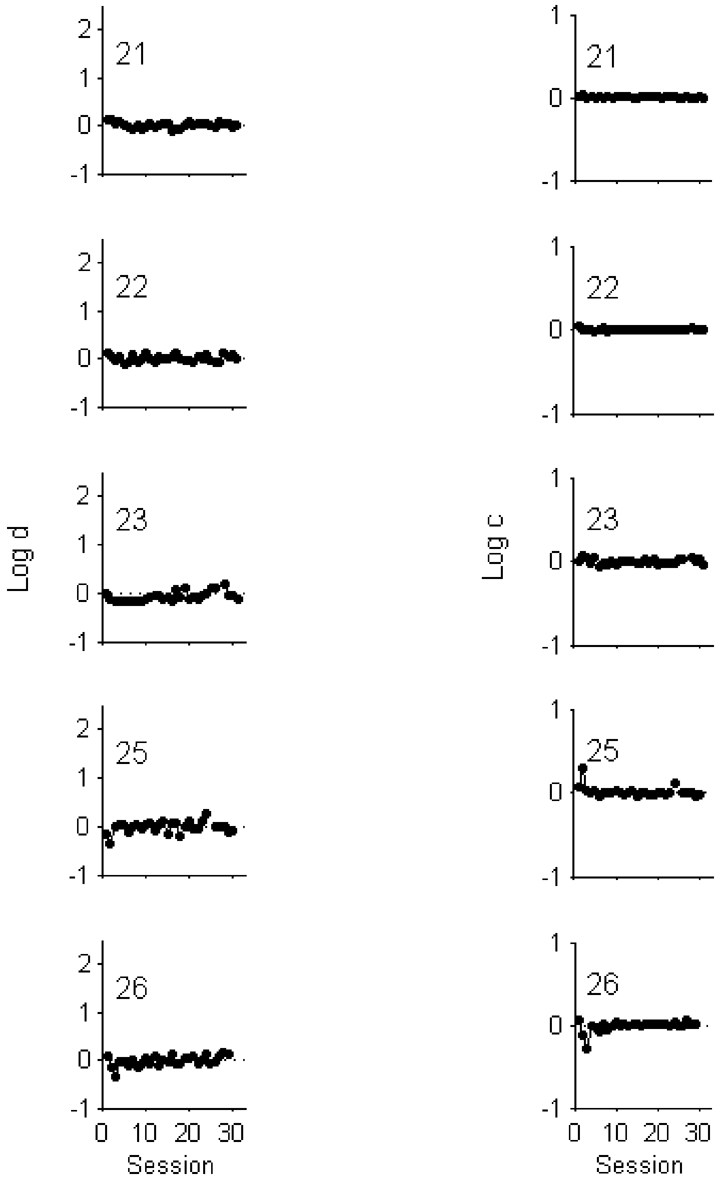


Figure 3.5b. The left panel shows the log *d* estimates for all five hens plotted against session number for Condition 2. The right panel shows the log *c* estimates plotted against session number.

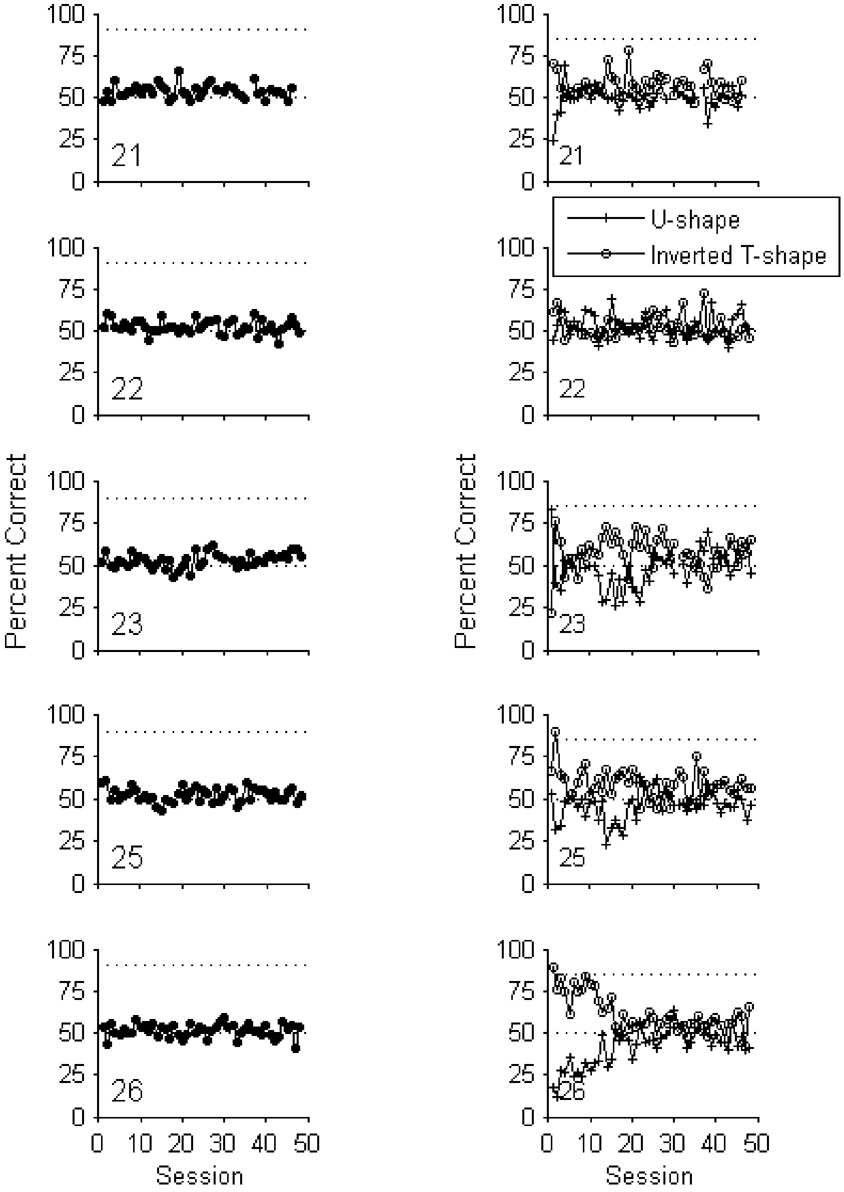


Figure 3.6a. The left panel shows the overall percentage correct with the blue and white Lego™ blocks (Condition 3) for all five hens plotted against session number. The right panel shows percentage correct to each of the blue and white Lego™ stimuli. The horizontal dotted lines mark 50 and 85 %.

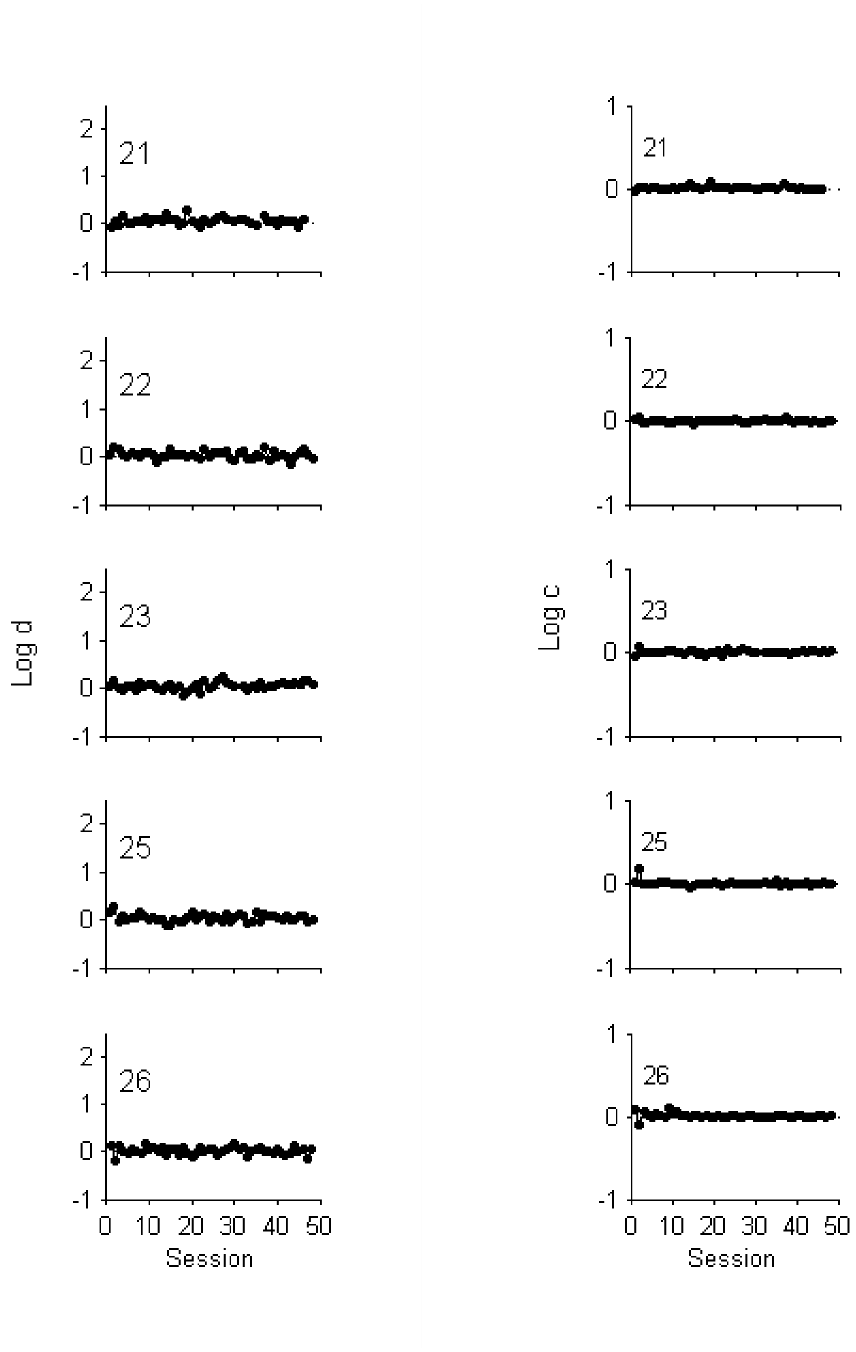


Figure 3.6b. The left panel shows the log *d* estimates for all five hens plotted against session number (Condition 3). The right panel shows log *c* estimates plotted against session number.

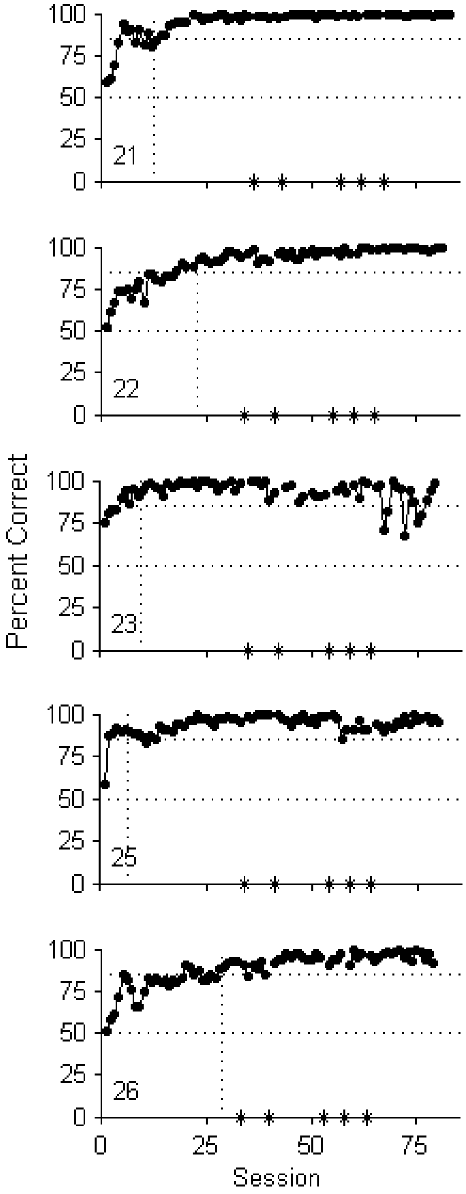


Figure 3.7a. Overall percentage correct over the training sessions with the colour stimuli (Condition 4) plotted against session number for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 and 85 %. The asterisks on the x axis represent the locations of the test sessions.

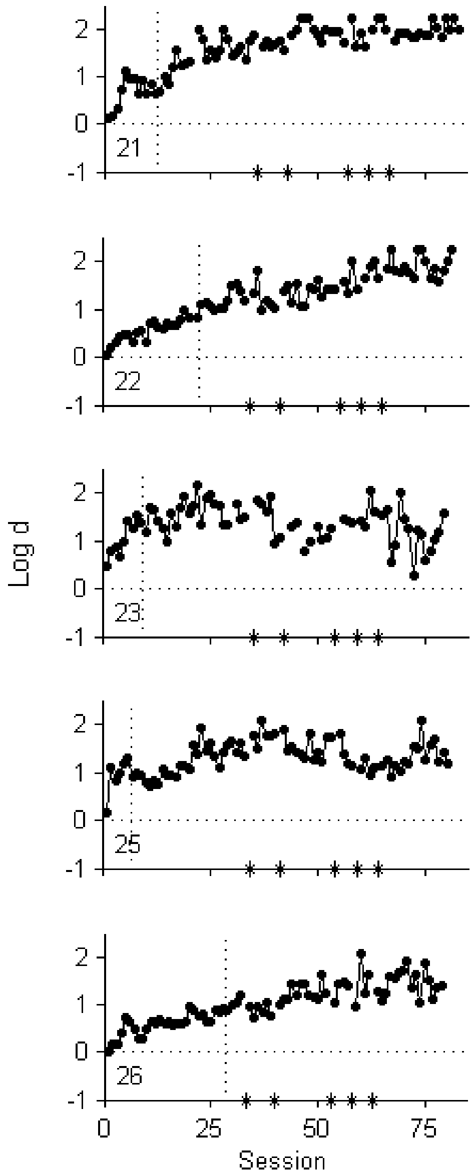


Figure 3.7b. Log *d* estimates over the training sessions with the colour stimuli (Condition 4) plotted against session number for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

Figures 3.8a and 3.8b present the percentage correct for each stimulus in Condition 4 and the log *c* estimates across sessions. There were no consistent differences in accuracy across the two stimuli, except Hen 23's performance was more variable with the blue stimulus from session 70 onwards. For all hens, there were also no consistent response biases for one key over another.

Figure 3.9 shows the mean percentages correct and standard deviations over the last five training sessions with the TFT monitor for each hen. Also plotted are the data from the test sessions with the CRT monitor set at the different frequencies for all five hens. The data from the replication of 100 Hz is shown as an x. The bottom right panel shows the percentage correct and standard error of the means at each frequency averaged across all five hens. Hen 22's accuracy remained high across all of the test sessions. However, accuracy generally decreased for all other hens as the refresh rate of the CRT monitor was decreased. On average, accuracy decreased and variability across hens increased as the refresh rate of the CRT monitor was decreased. There was no significant difference between the data from the first and last 100 Hz test sessions (paired-samples *t*-test, $t(4) = -0.292$, $p > .05$, $d = -.139$ (moderate effect size)). Hence, the averages of the 100 Hz test conditions were used for all further analyses. A repeated-measures ANOVA over the data from the test sessions showed that these changes were statistically significant ($F(4,16) = 5.072$, $p < .05$, $\eta^2 = 0.559$ (moderate effect size)). However, only Hen 23's data revealed a significant trend (Theil test, C = 8, $n = 5$, $p < .05$). None of the other hens performances showed a significant trend (Theil test, C = 4, 2, 6 and 6, for Hens 21, 22, 25 and 26 respectively, $n = 5$, $p > .05$ in all cases,). Further analyses with paired-sample *t*-tests showed that accuracy when the CRT monitor was set at 100, 85, and 75 Hz was not significantly different from that with the TFT monitor ($t(4) = 1.020$, $p > .05$, $d = .486$; $t(4) = 2.161$, $p > .05$, $d = .999$; $t(4) = 1.913$, $p > .05$, $d = 1.067$ (moderate effect size) respectively). Accuracy when the CRT monitor was set at 60 Hz was significantly different from accuracy in training with the TFT monitor ($t(4) = 3.403$, $p = < .05$, $d = 1.818$ (moderate effect size)). When the CRT monitor was set at 100 Hz there was no significant difference in accuracy from accuracy with the CRT monitor set at either 85 or 75 Hz ($t(4) = -.063$, $p > .05$, $d = -.027$ (no effect); $t(4) = 2.411$, $p > .05$, $d = 2.563$ (moderate effect size) respectively), there was, however, a significant difference between accuracy with the CRT monitor set at 100 Hz and at 60 Hz ($t(4) = 4.611$, $p < .05$, $d = 2.565$ (moderate effect size)). There was no significant

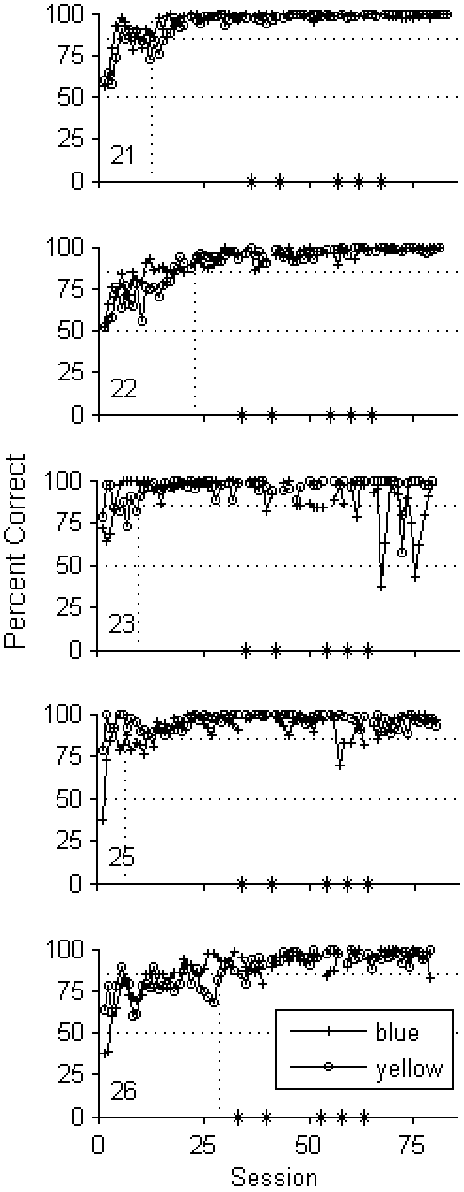


Figure 3.8a. Percentage correct for each of the colour stimuli (Condition 4) over training sessions plotted against session number for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 and 85 %. The asterisks on the x axis represent the locations of the test sessions.

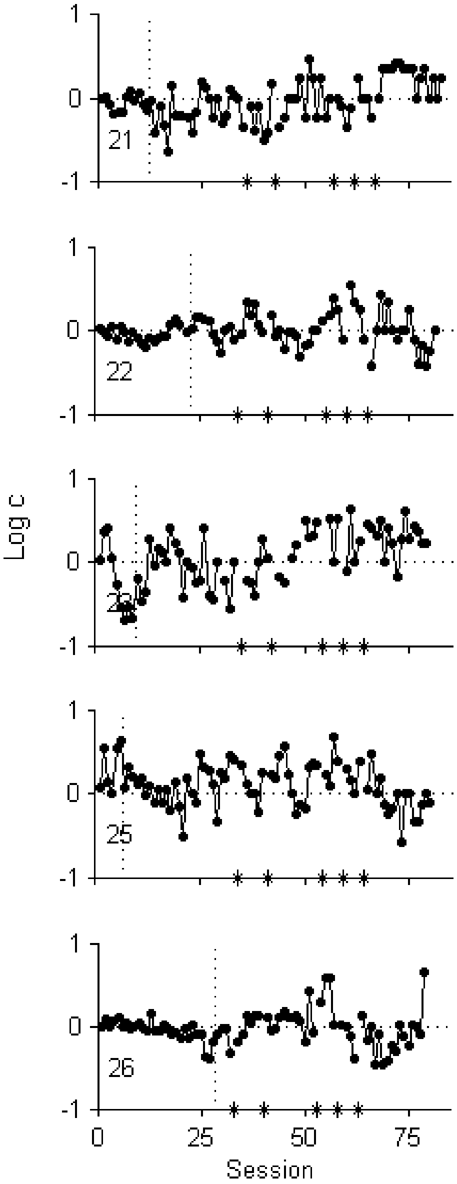


Figure 3.8b. Log *c* estimates plotted against session number for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

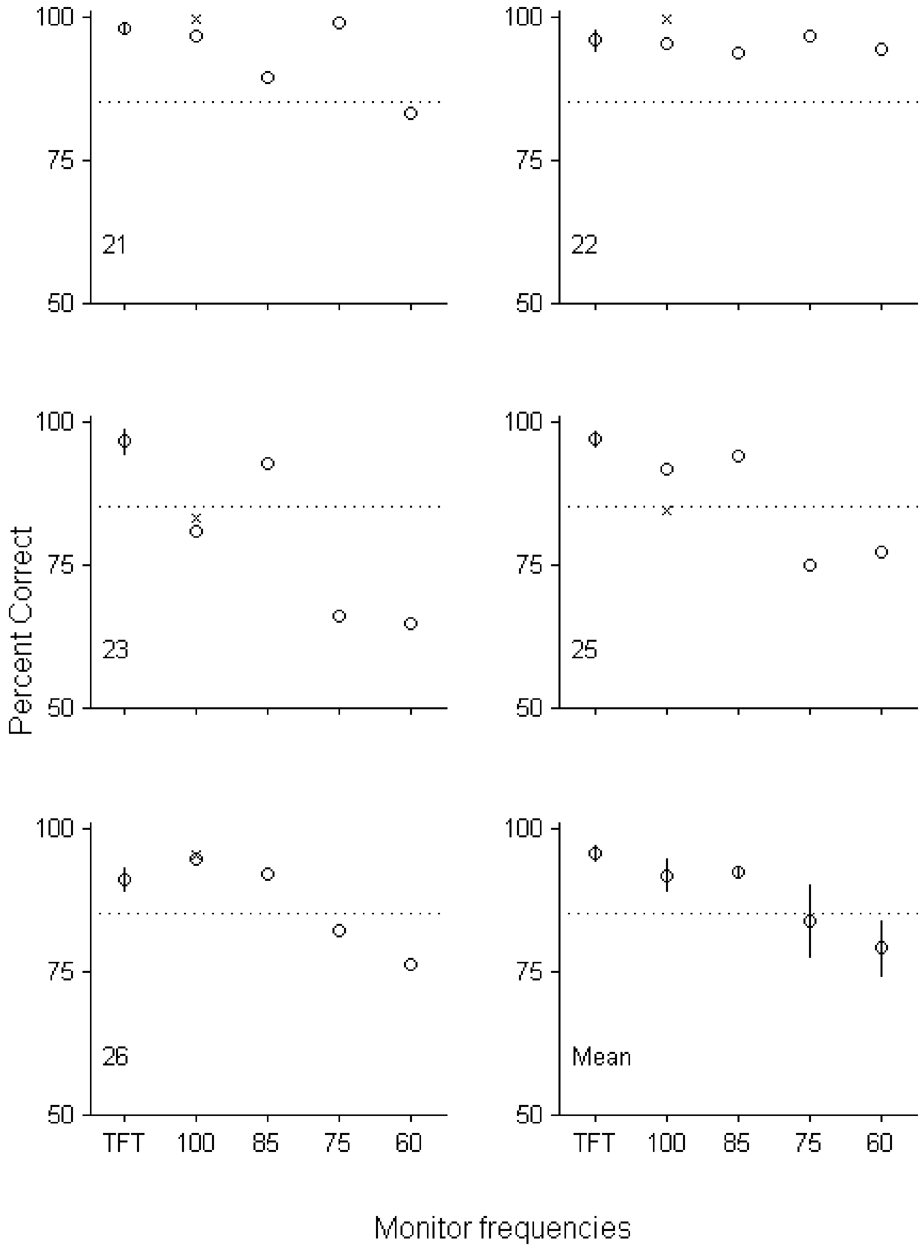


Figure 3.9. Mean percentages correct and standard deviation over the last five training sessions (TFT) and the data from the test sessions with the CRT monitor for the colour stimuli (Condition 4). The vertical lines mark one standard deviation each side of the mean and the dotted horizontal line shows 85 %. The replication of the 100 Hz condition is shown as x. The bottom right panel shows the average percentage correct and standard error of the means at each frequency level across all five hens.

difference in accuracy when the CRT monitor was set at 85 Hz from accuracy at 75 or 60 Hz (paired-samples *t*-test, $t(4)=1.286$, $p>.05$, $d=.631$; $t(4)=2.706$, $p>.05$, $d=1.497$ (moderate effect size)), and there was no difference between accuracy with the CRT monitor set at 75 and 60 Hz ($t(4)=1.526$, $p>.05$, $d=.780$).

Percentages correct for blocks of 10 trials were examined for each of the 25 test sessions for this condition. In five of these test sessions there was an increasing trend in accuracy (Hens 23 and 25 in the first 100 Hz test session, and Hens 21, 25 and 26 in the 60 Hz test session). There was no visual trend in accuracy in all other test sessions.

The left panel of Figure 3.10a shows the overall percentages correct for the pictures of the blue and white Lego^{TM} blocks with the sides removed (Condition 5) for all sessions. The left panel of Figure 3.10b shows the log *d* estimates across sessions. None of the hens' accuracies was above 85 % over the 80 training sessions, but two data sets (Hen 22 and 26) trend upwards from about session 20, but both then stabilised below 85 % correct from session 50. On average, the total number of trials per session ranged from 100 to 180 across hens. For 100 trials per session, any percentage correct over 60 % would be significantly different from chance at a .05 level of significance (binomial test). Thus, Hens 22 and 26's percentages correct were significantly above chance, averaging 84.6 and 75.8 %, respectively, over the last five training sessions. The remaining three hens responded at around chance levels throughout these training sessions. Hen 23 did not respond at all for a number of sessions as shown by the number of missing data points. The right panel of Figures 3.10a and 3.10b presents the percentage correct for each stimulus in Condition 5, and the log *c* estimates, across sessions for all hens. In general, all hens showed higher accuracy with the inverted T-shape than with the U-shape, however, only Hens 22 and 26 showed a bias towards selecting the right key over the left key.

The left panel of Figures 3.11a and 3.11b shows the overall percentage correct and the log *d* estimates over all training sessions for the line drawing discrimination (Condition 6) for all hens. All hens started at around 50 % correct but showed an overall increase in discrimination as shown by the increasing data paths. After 25 to 31 training sessions, Hens 22, 23 and 26 had completed five sessions at or above 85 %. Hen 21's accuracy trended upward and reached 85 % for five sessions after 44 sessions and during the period in which the test sessions were conducted. Hen 25's accuracy was below 85 % over all of the 75 training sessions, however, her

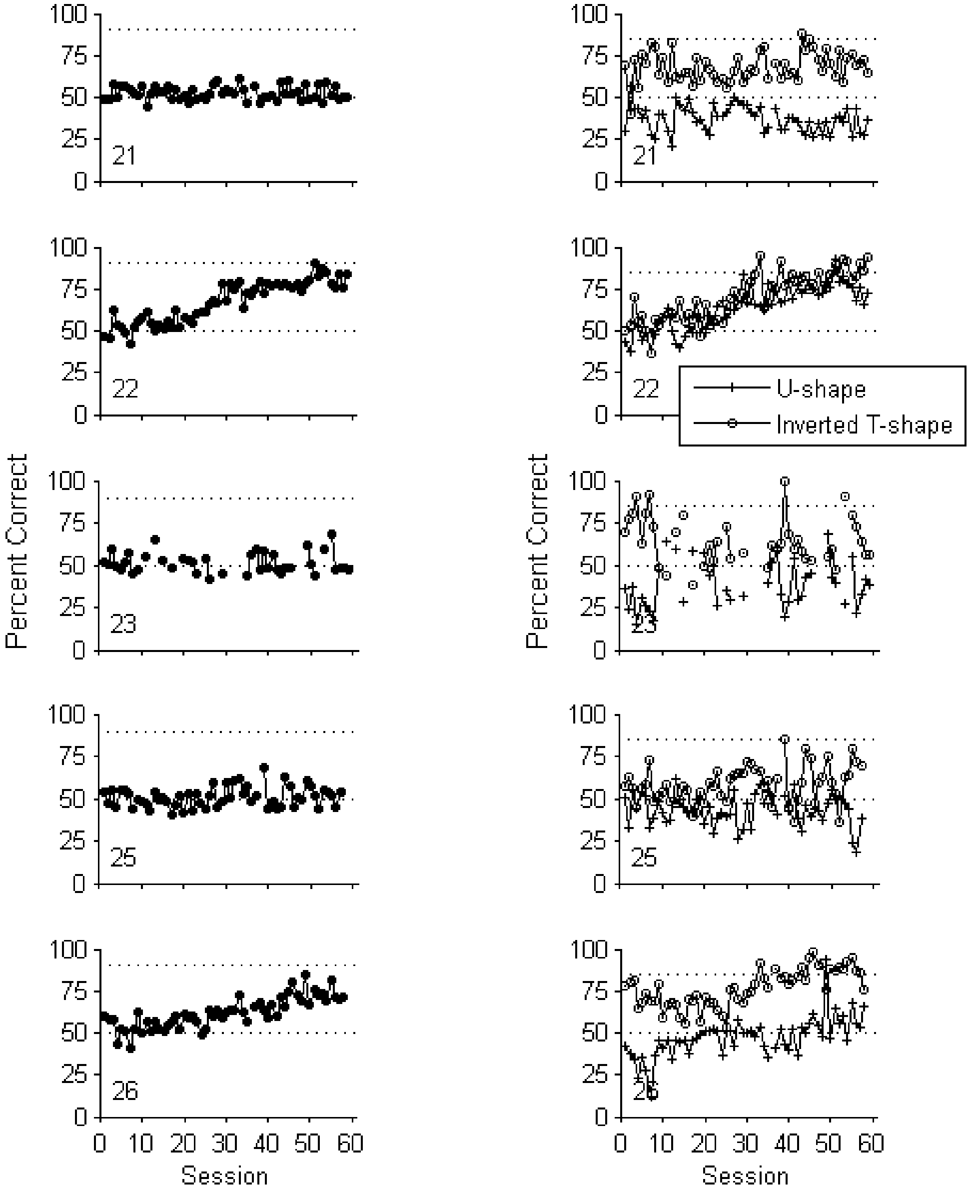


Figure 3.10a. The left panel shows the overall percentage correct with the blue and white Lego™ blocks (Condition 5) for all five hens. The right panel shows the percentage correct for each blue and white Lego™ block stimuli. The horizontal dotted lines mark 50 and 85 %.

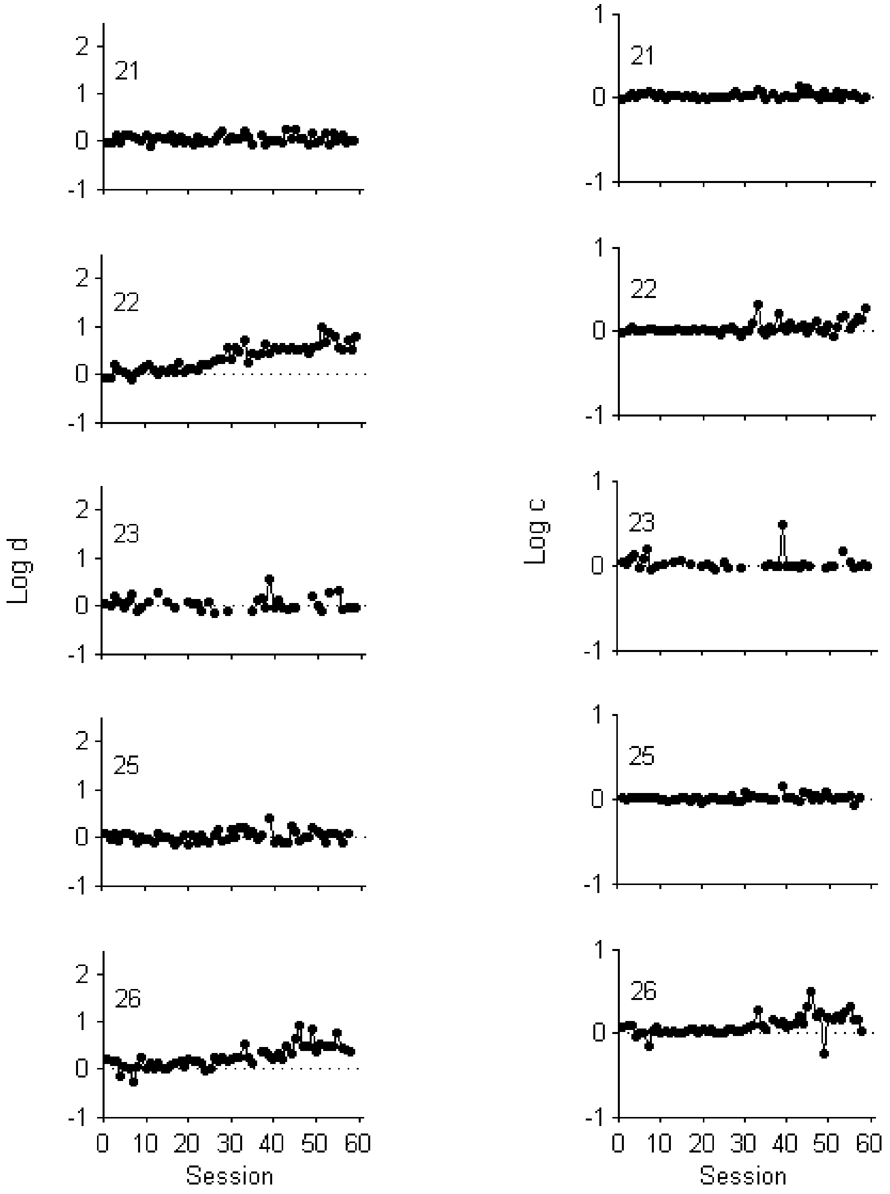


Figure 3.10b. The left panel shows the log *d* estimates for all five hens across session number. The right panel shows the log *c* estimates across session number.

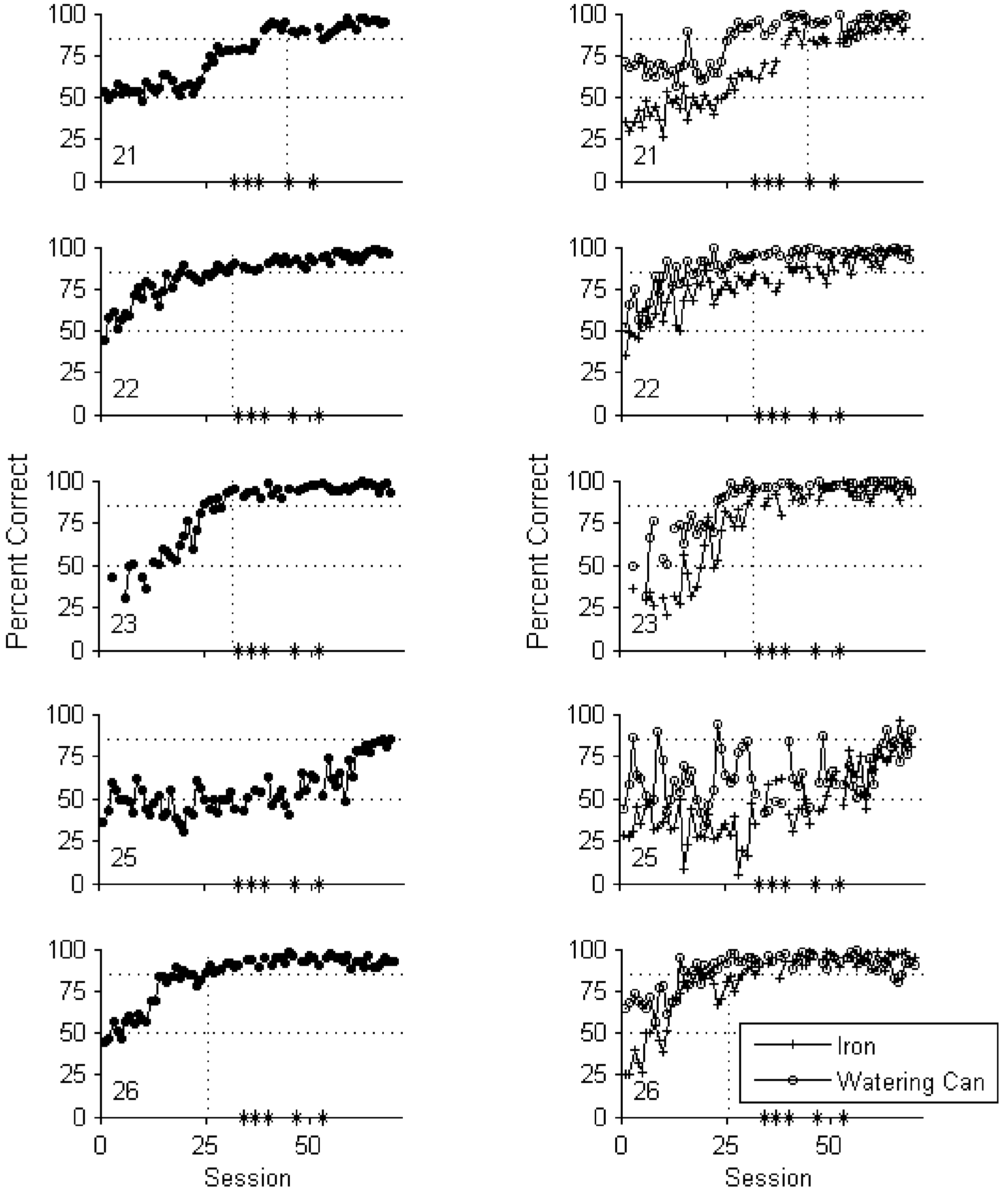


Figure 3.11a. The left panel shows the overall percentage correct over the training sessions with the line drawing stimuli (Condition 6) plotted against session number. The right panel shows the percentage correct for each of the line drawing stimuli. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 and 85 %. The asterisks on the x axis represent the locations of the test sessions.

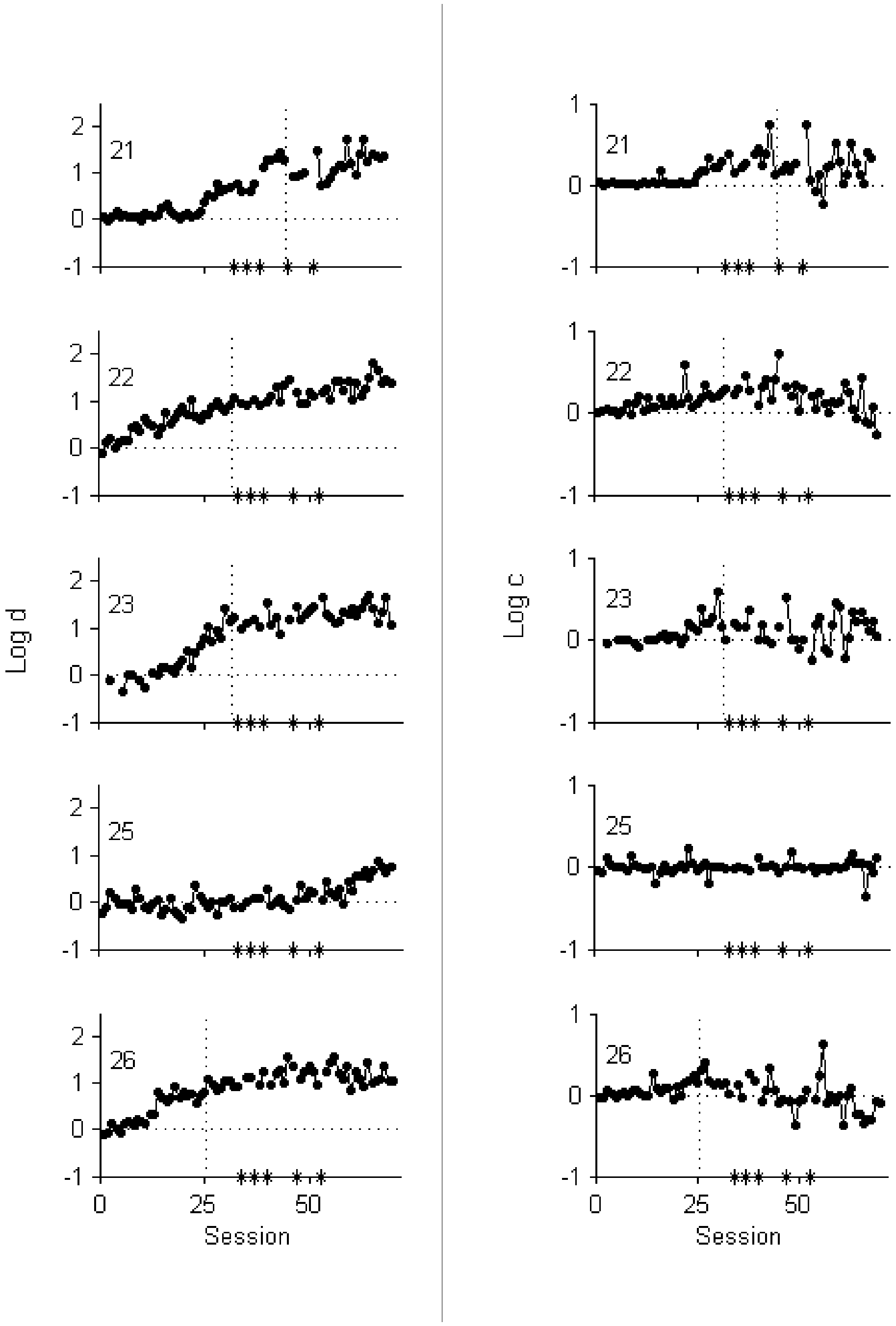


Figure 3.11b. The left panel shows the log *d* estimates plotted against session number. The right panel shows the log *c* estimates plotted against session number. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

accuracy trended upwards over the later sessions. Between every test session each hen had a minimum of two training sessions with the TFT monitor. Accuracy for these interspersed sessions ranged from 78 to 98 %.

The right panel of Figures 3.11a and 3.11 b presents percentage correct for each stimulus and the log c estimates across sessions in Condition 6. During the early training sessions, all hens showed slightly higher accuracy with the line drawing of the watering can rather than the line drawing of the iron. There were, however, no large differences in accuracy with the two stimuli over the final training sessions. Over the later training sessions, Hens 21, 22 and 23 showed a slight bias towards responding more on the right key over the left key. Hens 25 and 26 showed no consistent biases in responding to either key.

Figure 3.12 shows the mean percentages correct and standard deviations over the last five training sessions with the TFT monitor and percentage correct from the test sessions. The data from the replication with the 100 Hz is shown as an x. The bottom right panel shows the percentage correct and standard mean error at each frequency level averaged over the four hens whose accuracies were at or above 85 % during training. There was no significant difference between the data from the first and last 100 Hz test sessions (paired-samples t-test, $t(3) = -1.306$, $p > .05$, $d = -.663$). Hence, the averages of the data from the 100 Hz test sessions were used for all further analyses. Generally, accuracy decreased and variability increased as the CRT refresh rate decreased, although accuracy at 75 Hz was slightly higher than that at 100 Hz or 85 Hz.

A repeated-measures ANOVA over the data from the test sessions showed that the overall changes were statistically significant ($F(4,12)=11.833$, $p<.05$, $\eta^2=0.798$ (strong effect size)). However, Hens 23 and 26's data showed a significant trend (Theil test, C = 7 and 8 respectively, $n= 5$, $p<.05$). Hens 21 and 22's data showed no trend (Theil test, C = -1 and 6 respectively, $n= 5$, $p>.05$)

Accuracy when the CRT monitor was set at 100 Hz was not significantly different from that with the TFT monitor (paired-sample t-test, $t(3)=2.559$, $p>.05$, $d= -1.306$ (moderate effect size)). However, paired-samples t-test showed that accuracy with the CRT monitor set at 85, 75, and 60 Hz was significantly different from that with the TFT monitor ($t(3)=4.727$, $p<.05$, $d=2.446$ (moderate effect size); $t(3)=5.683$, $p<.05$, $d=2.868$ (strong effect size); $t(3)=4.264$, $p<.05$, $d=2.774$ (strong effect size)

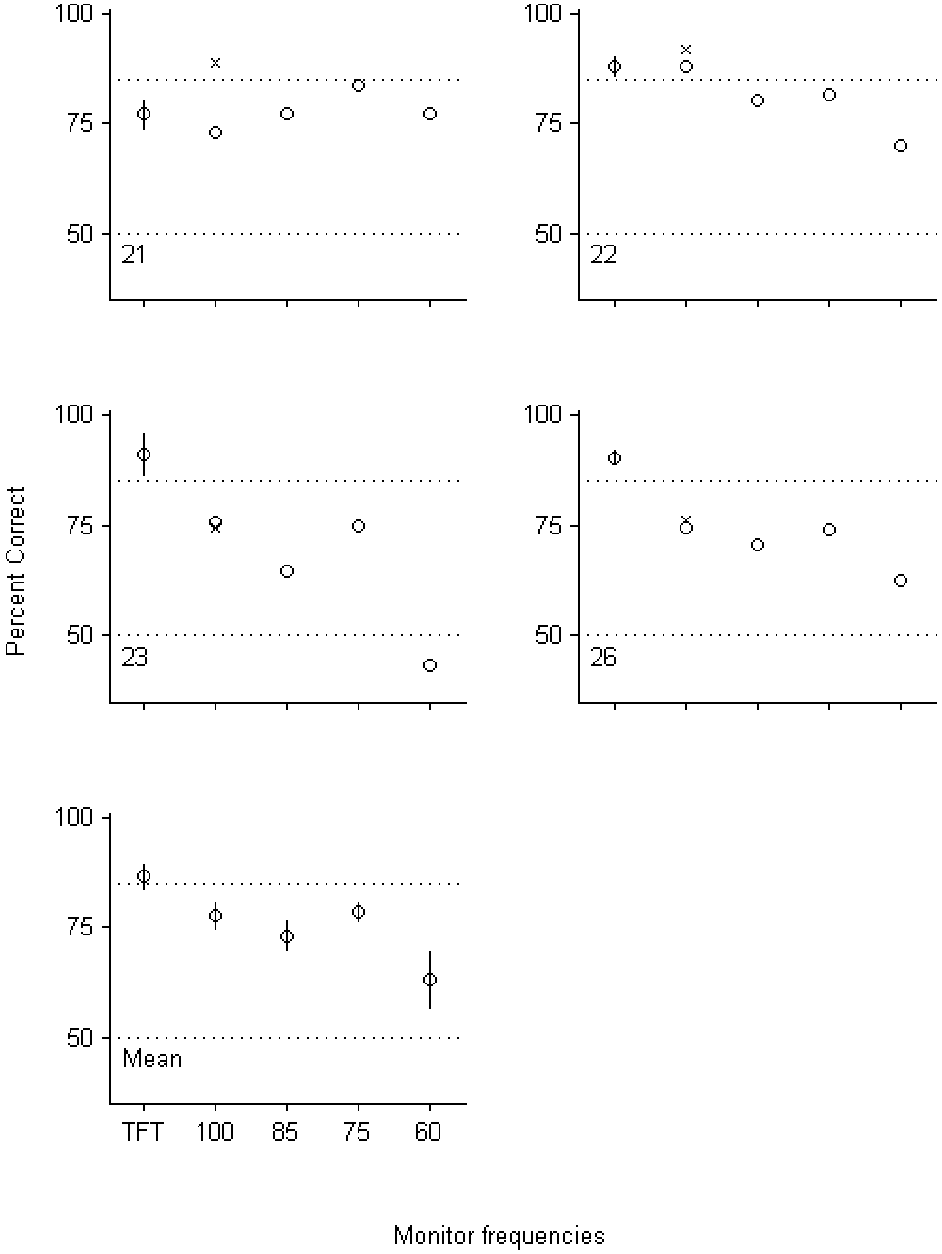


Figure 3.12. Mean percentages correct and standard deviation over the last five training sessions with the TFT monitor (TFT). Also plotted are the data from the test sessions with the CRT monitor for the line drawing stimuli (Condition 6) for those hens that had at least five sessions at or above 85 %. The vertical lines mark one standard deviation each side of the mean and the dotted horizontal line shows 85 % accuracy levels. The replication of the 100 Hz condition is shown as x. The bottom left panel shows the average percentage correct and standard error of the means at each frequency level.

respectively). Accuracy with the CRT monitor set at 100 Hz was significantly different from that with the CRT monitor set at 85 Hz (*t*(3)=4.158, *p*<.05, *d*=2.078 (moderate effect size)), but was not significantly different from that with the CRT monitor set at 75 or 60 Hz (*t*(3)=0.762, *p*>.05, *d*=.406; *t*(3)=2.861, *p*>.05, *d*=1.750 (moderate effect size) respectively). There was also no significant difference in accuracy when the CRT monitor was set at 85 Hz from accuracy with it set at 75 or 60 Hz (*t*(3)=-2.723, *p*>.05, *d*=-1.586 (moderate effect size); *t*(3)=2.263, *p*>.05, *d*=2.086 (moderate effect size); respectively), or between the CRT monitor set at 75 and 60 Hz (*t*(3)=2.766, *p*>.05, *d*=-2.481 (moderate effect size)).

Percentages correct for blocks of 10 trials were examined for each of the 25 test sessions for this condition. Five of the test sessions showed small but increasing trends in accuracy (Hen 22 in the 85 Hz test session, Hen 25 in the 75 Hz test session, Hen 21 in the 60 Hz session, and Hens 21 and 26 in the second 100 Hz test session). There were no visual trends in accuracy for the other twenty test sessions.

Discussion

The aim of this experiment was to investigate how altering the refresh rate of a CRT monitor would affect hens' accuracy during the transfer-test sessions. Most of the hens learned to discriminate between shape stimuli (cross and circle), coloured stimuli (blue and yellow squares), and the line drawings (watering can and iron). When the hens had been able to discriminate the stimuli shown on the TFT monitor, accuracy was generally high in the test sessions with the CRT monitor set above the CFF of the hens, and was lower when the CRT monitor was below their CFF (see Figure 3.13). Hence, decreasing the refresh rate of the CRT monitor decreased the degree of transfer of the previously learned discrimination. That is, when the discrimination was learned with 'steady' images, the flickering images disrupted performance. As the refresh rate decreased, accuracy decreased, suggesting the stimulus appeared increasingly different from the original stimulus. This finding suggests that the previously reported failures of bird species, especially hens, to transfer from real images to stimuli on CRT monitors may have been because the stimuli appeared to be flickering on CRT monitor.

Although it would have possibly been desirable to have counterbalanced the order of the test sessions, the effects of each test session could be seen in the performance in the training sessions between each test session. All hens' accuracies

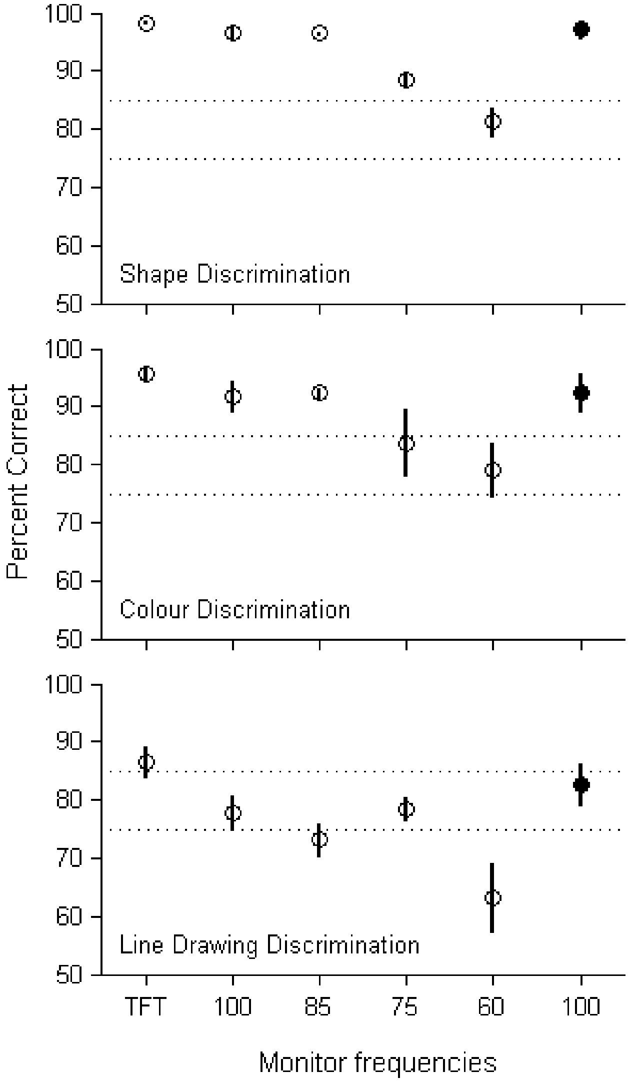


Figure 3.13. Mean percentages correct and standard error of the means over the last five training sessions with the TFT monitor (TFT) and the data from the test sessions with the CRT monitor set at the different frequencies. The vertical lines mark one standard deviation each side of the mean. The replication of the 100 Hz condition is shown as a filled in circle. The dashed horizontal lines show 75 % and 85 % accuracy levels. The top panel represents the data from the shape discrimination (Condition 1). The second panel represents the data from the colour discrimination (Condition 3). The third panel represents the data from the line drawing discrimination (Condition 6).

recovered fully in these training sessions showing that there was probably no effect of the order of the test sessions. In addition, the data from the final 100 Hz test session were comparable to those from the first 100 Hz test session for all stimulus pairs, suggesting there were no cumulative sequential effects of the test session order.

Although the hens learned the colour, shape and line drawing discriminations, none of them learned to discriminate between the various images of the Lego^{TM} stimuli (Conditions 2 and 3). While the stimuli appeared to differ in appearance to the human eye, it may be that the stimuli were too similar in appearance for the hens to distinguish. Yet, previous studies have shown hens can learn to discriminate a range of stimuli such as gratings and grey stimuli (DeMello, Foster & Temple, 1992, 1993), colours (Nakagawa, Etheredge, Foster, Sumpter & Temple, 2004; Poling, Temple, & Foster, 1996), geometric stimuli (Werner & Rehkämper, 1999; also shown in Experiment 2), line orientation (Werner, Tiemann, Cnotka & Rehkämper, 2005) and pictures of conspecifics (Candland, 1969; Weavers, 2000). It seems unlikely that the hens were unable to see differences across the stimuli in Conditions 2 and 3, as DeMello et al. (1992) showed that hens have high visual acuity. Thus, it is not clear why the hens were unable to learn these particular discriminations.

In the present experiment, although the hens' performances remained around 50 % correct over all training sessions, they still continued to respond. Learning could possibly have been encouraged during these conditions in a number of ways. For example, the schedule of reinforcement could have been made leaner by increasing the variable ratio schedule. The VR schedule in the present experiment started rich, because when too few reinforcers are gained during a session, animals tend to stop responding when learning a new task. To maintain an equal number of reinforcers across sessions, and to allow an increase in the number of trials per session, the ratio was increased as the hen's accuracy and rate of responding increased and as, in so doing, the hen gained more of the reinforcers. In most conditions the VR was increased to 5. However, in Conditions 2 and 3, the hens' accuracy remained around 50 %, and they were only gaining, on average, 13-16 reinforcers per session, and so the ratio was held at 4. It could be possible that increasing the VR schedule might have increased the hens' accuracy. But although responding in a session seemed to be affected by reinforcement rate, no published data could be found suggesting that this might be the case. Further research is required into how altering reinforcement rate may affect behaviour.

Another possible method to increase the hens' accuracy could have been to use a correction procedure. In such a procedure incorrect trials are repeated until the animal makes a correct response. Correction procedures can be useful in that they may reduce key or stimulus biases (Mackay, 1991 cited in Iversen, 1993). Without correction procedures, biases may still be maintained through intermittent reinforcement, and accuracy may remain low. Thus, accuracy may increase if correction procedures are used. However, the hens in the present experiment did learn to discriminate between some of the stimuli without the inclusion of a correction procedure. This indicates that the procedure used here was sufficient for the hens to learn some of the conditional discriminations.

One possible reason for the failure of the hens to learn the discrimination in Conditions 2, 3 and 5 may have been the three-dimensional cues (e.g., shading or object sides) that were present in the images may have made it difficult for the hens to see the outlines of the objects. These images represented objects that clearly differed in shape to the human eye. The images were even more clearly different to the human eye when these features were removed from the images. Even with this change four of the hens did not learn the discrimination (Condition 5). These Lego shapes differed only in the shape of their top sections. These same hens were able to discriminate between the cross and circle stimuli and between the two line drawings. These other stimuli differed in more than one features and so had more differences than the Lego shapes on which the hens could base discrimination. This seems to have made these discriminations easier.

The colour stimuli used in Condition 4 were a blue or yellow square on a white background, and the hens' discriminative performance decreased over test flicker frequency. These results are in contrast to those found in Experiment 2, where the performance of the hens remained high when tested with the colour stimuli (red and green square on a black background) on the CRT monitor at 60 Hz. One possibility is that hens find it more difficult to discriminate between the colours yellow and blue, than they do red and green. However, this seems unlikely as outlined above, hens have been shown to respond differently to blue, yellow, red and green stimuli (Huber-Eicher, 2004; Jones & Carmichael, 1998; and Zupan, Kruschwitz, & Huber-Eicher, 2007) showing that hens are able to distinguish between all of these colours.

Another possibility for the poor performance in Condition 4 is the size of the stimuli. The coloured stimuli used in Experiment 2 were more than twice as large (108-mm wide × 108-mm high) as those used in Experiment 3 (45-mm wide × 47-mm high). Thus a possible reason for the different results is that the edges of the smaller stimuli were more obvious, and the flicker was more apparent, thereby disrupting behaviour. Alternatively, the differences in the size of the stimuli may have made the flicker more apparent to the hens.

The main advantage of the method used in present experiment, in comparison to that used in Experiment 2, was that the observing response could be used in testing sessions. The average percentages correct using the shape stimuli (cross and circle) were calculated for each test session used in both Experiments 2 and 3 and these are presented in Table 3.2. This was done to determine if the inclusion of the FR requirement during the test sessions had an effect on the hens' performances. The hens were more accurate when the FR requirement was included. This finding implies that the FR requirement served to not only improve the hens' accuracy during the training sessions, but also to improve accuracy during the test sessions. However, this increase in accuracy may also have been a result of the hens' previous experience with these particular stimuli in the previous experiment.

A way of viewing the present finding that accuracy decreased as the refresh rate was decreased comes from comparison with the study of stimulus control over behaviour. Typical stimulus control studies (see for example, Ghirlanda & Enquist, 2003; Guttman & Kalish, 1956; Honig & Urcuioli, 1981) examine the effects of changes in one dimension of the training stimuli (e.g., wavelength) on animals' performances. The resulting functions generally show decreases in response rate with increasing changes in the stimuli. The data in Figure 3.13 follow this pattern showing decreased accuracy with decreasing refresh rates. However, stimulus control studies do not normally involve conditional discriminations, as used here, and generally the functions plot the degree of change in response rate rather than the change in accuracy. The degree to which the response rate decreases is said to reflect the level of stimulus control.

Although response rate data are typically used in generalisation gradients there are some studies that use other measures. For example, Sargisson and White (2001) suggested that proportions correct could be interpreted in a similar way. They trained groups of pigeons in a delayed matching-to-sample (DMTS) task with

Table 3.2.

The average percentages correct across all hens in each of the test sessions using the shape stimuli for Experiments 2 and 3.

Experiment	100 Hz	75 Hz	60 Hz
2	85.3	69.3	62.8
3	95.3	82.4	81.3

particular delays, and then tested the pigeons using delays around the training delay. They found decreases in proportion correct when delays longer than the training delay were presented. They argue this showed stimulus control generalised to some degree to delays close to the training delay. In the present study, as in Sargisson and White (2001), the task was a conditional discrimination, and percentages correct, rather than response rates, were examined. Percentage correct generally decreased with decreasing refresh rates, showing that the stimulus control decreased as refresh rate decreased. Accuracy on the CRT monitor at 100 Hz for the colour and shape stimuli was similar to that found with the TFT monitor. This shows an almost equivalent degree of control and suggests that the images on both monitors appeared similar to the hens. Interpreted in this way, the decrease in accuracy with decreasing flicker suggests that the flickering stimuli are similar but not identical to the TFT images and the greater the flicker the less similar they are.

The present results suggest that images presented on CRT monitors that are refreshed at rates lower than the CFF of an animal will not appear the same as the real object. If images are to be used as a substitution for a real stimulus then TFT monitors, or CRT monitors set at a refresh rate above the animal's CFF, should be used. It must be noted, however, that the performance of the hens in this study never fell below an average of 63 % correct, and was often above 75 % correct. Thus, some degree of discrimination remained even at low refresh rates and some properties of the original stimulus must have still been visible for the hens to be able to make some discrimination. The degree to which the discrimination broke down was similar across the colour stimuli and the shape stimuli. The line drawings resulted in a steeper drop in accuracy with decreasing flicker than did the shape or colour stimuli. Accuracy with the CRT monitor set at 100 Hz was lower than with the TFT monitor and it appears even this degree of flicker changes the appearance of the line drawings. The line drawings had more points of difference between the two stimuli than the colours or shapes, and it may be that these differences are more easily disrupted when the stimuli appear flickering. Thus, these results may indicate that discrimination between the line drawings may be more disrupted by changes in refresh rate.

It seems reasonable to suggest that the stimuli appeared altered when the refresh rate decreased and that they may be more difficult for the hens to discriminate when they are presented at a flicker frequency that is below their CFF. However, the decrement in performance may simply be because the stimulus looks increasingly different from the training stimulus, thus changing the discrimination. It is clear that decreasing flicker decreased accuracy when the original discrimination was a steady image. This could be a result of the images becoming more difficult for the hens to see in the same way that the increasing the number of gratings on a grating stimulus makes it harder to discriminate from a grey stimulus. That is, it is the animal's visual system that gives rise to the difficulty to discriminate between the stimuli. Alternatively, the flicker may change the appearance of the stimuli, so that test sessions are rather like presenting a new discrimination. The next study was designed to see if the hens could learn to discriminate stimuli presented on a CRT monitor at 60 Hz, and whether they could transfer this discrimination to stimuli that appeared either less flickering or 'steady'.

EXPERIMENT 4

In the previous experiment, hens were trained to discriminate stimuli presented on a TFT monitor and then tested for transfer of the discrimination to a CRT monitor set at various refresh rates. It was found that accuracy generally decreased as the monitor refresh rate decreased. There are at least two possible explanations for this finding. First, it could be that the low refresh rates made the images less discriminable. That is, the discrimination was the same, but it was made more difficult by the change in flicker rate, and hens' visual system may be limited in its ability to discriminate stimuli presented at close to or below their CFF. Second, when the images are presented at a low flicker rate then they may not be immediately recognisable as the original training stimuli. That is, the flickering stimuli might appear to be different from the training stimuli and thus the discrimination task during the test sessions was a novel one.

One way to investigate this further would be to train hens in a discrimination presented on a CRT monitor set at 60 Hz (i.e., below their CFF and assumed to be seen as flickering by the hens). Then, if the hens learn this discrimination, transfer tests could be conducted using the same images but now presented on a CRT monitor set at higher refresh rates.

One possible outcome of such a procedure would be that the hens do not learn the discrimination at low flicker rates. If so, this would suggest that the stimuli are hard to discriminate when the monitor is set at 60 Hz. Another possible outcome would be that the hens do learn to discriminate the flickering stimuli, but then not transfer this learning to the same stimuli presented on the monitor set at higher refresh rates. This would suggest that hens can discriminate flickering images, but that they are not equivalent to images presented at a higher refresh rate (i.e., steady images). A further possible outcome would be that hens may learn to discriminate the flickering stimuli, and they may also show transfer to the same stimuli presented using higher refresh rates. This would suggest that once they have learned to discriminate flickering images, they can transfer this to the steady, or faster flickering, images. Within this range, it is possible that once the discrimination is learned, different degrees of transfer may be seen, showing some degree of stimulus control resulting from the flicker of the CRT monitor.

The next study used this procedure. Initially the two stimuli were presented on a CRT monitor set at 60 Hz and, using the same procedure as in Experiment 3, the hens were required to learn the discrimination. This was followed by testing for the transfer of this discrimination to the images presented at higher refresh rates and on a TFT monitor. To compare the findings with those of the previous experiment, similar stimuli were needed and so a pair of line drawings, similar to those used in Experiment 3, were used as the stimuli. The drawings selected were two of those used by Wasserman et al. (1993) in a four-alternative forced-choice discrimination that assessed spatial organization in pigeons.

Method

Subjects

The five subjects were the same as those used in Experiment 3.

Apparatus

The apparatus was the same as that used in Experiment 3 except that the stimuli used were line drawings of a boat (45-mm high × 55-mm wide) and a lamp (60-mm high × 28-mm wide) (shown in Figure 4.1).

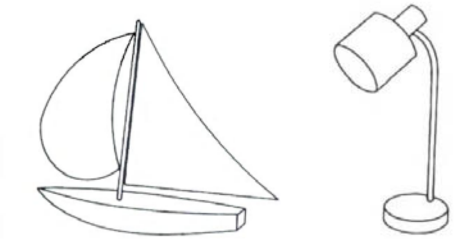


Figure 4.1. The line drawings of the boat and lamp stimuli.

Procedure

Discrimination Training. The procedure was the same as used in the discrimination training sessions in Experiment 3, except that the images were presented on a CRT screen set at 60 Hz. The correct response was to peck the left key if shown the boat, and peck the right key if shown the lamp. The VR requirement was gradually increased to VR 5 over 32 sessions. Test sessions began after at least two hens performances had reached five, not necessarily consecutive, sessions at or above 85 %.

Test Sessions. The test sessions followed the same procedure as used in Experiment 3. The stimuli were presented on the CRT monitor at higher refresh rates than those used during training and on a TFT monitor in the test sessions. The refresh rate of the CRT monitor was 75, 85 and 100 Hz over the first three test sessions. The stimuli were presented on the TFT monitor for the fourth test session, and the fifth test session was a repeat with the CRT monitor set at 75 Hz. Hen 23 did not respond during two of the test sessions; the first 75 Hz session and the TFT session. Each hen received at least two training sessions with the CRT monitor at 60 Hz after every test session.

Results

Figures 4.2a and 4.2b shows the overall percentage correct and the log *d* estimates over all training sessions. For each hen, the asterisks show the locations of the test sessions. Hens 21, 22 and 23's accuracy was at or above 85 % for five sessions, within 22, 43 and 68 training sessions respectively. The missing data points for Hen 23 are because this hen did not respond at all in some sessions, especially once the test sessions started. Hens 25 and 26 did not respond at 85 % (or above) correct for at least five sessions before the test sessions began, however, their accuracy increased over further training sessions. The minimum total trial numbers for these two hens was 185; therefore, any percentage correct above 57 % would be significantly different from chance (binomial test with α of .05). Hens 25 and 26 achieved averages of 77.8 % and 73.8 % correct, respectively, both significantly different from chance, over the last five sessions of training before the test sessions began. The log *d* estimates show that all hens' data paths show some increase in accuracy over sessions.

Figures 4.3a and 4.36b presents the percentage correct for each stimulus and the log *c* estimates across sessions for each hen. Over the later training sessions, there were no consistent differences in accuracy across the two stimuli Hens 22, 23, 24 and 25. Hen 26 was more accurate with the boat stimulus than with the lamp stimulus, and showed a bias towards selecting the left key. Alternatively, Hen 21 showed a bias towards selecting the right key.

Figure 4.4 shows the mean percentages correct and standard deviations over the last five training sessions before test sessions began (shown as o), and from five training sessions carried out after the second test (shown as □). It also shows the data

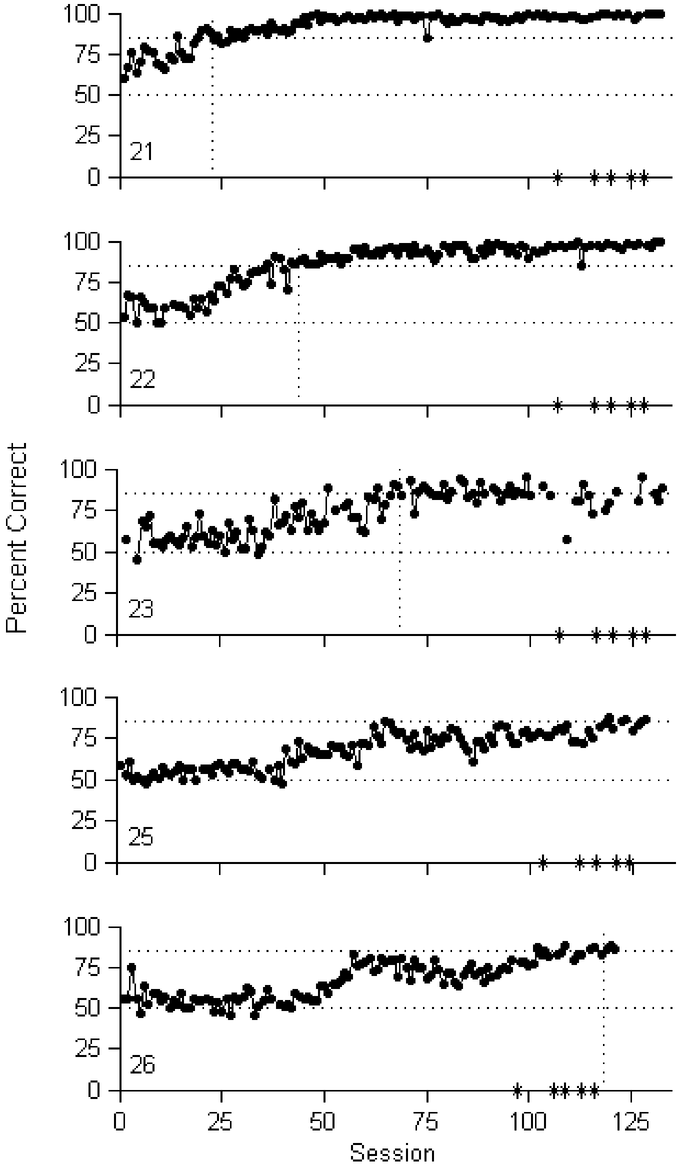


Figure 4.2a. Overall percentage correct plotted against session number for all hens. The horizontal lines mark 50 % and 85 %. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

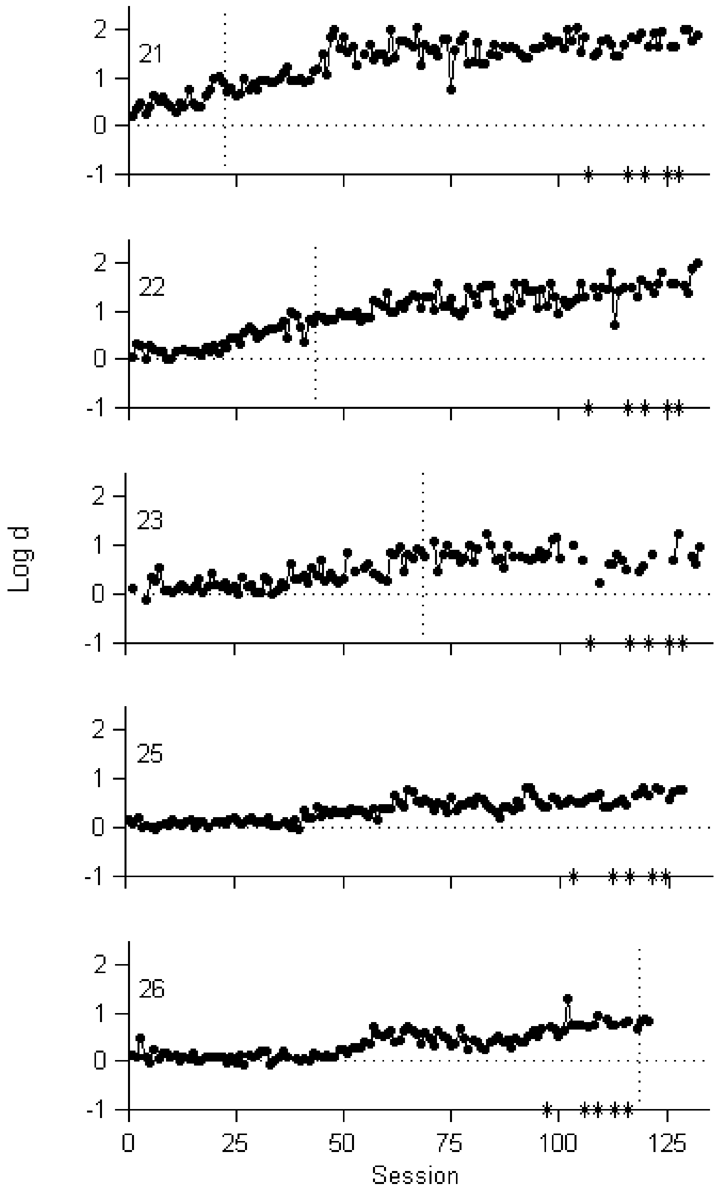


Figure 4.2b. Log *d* estimates plotted against session number for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

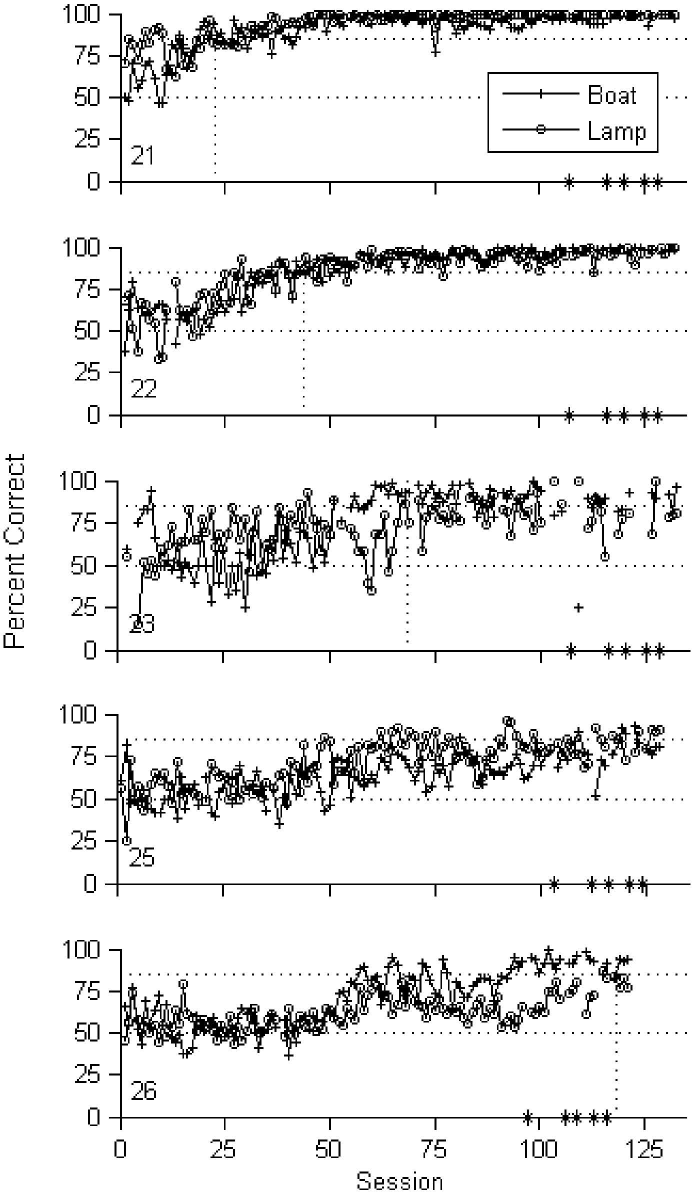


Figure 4.3a. Percentages correct for each of the line drawing stimuli over training sessions for all hens. The horizontal dotted lines mark 50 and 85 %. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

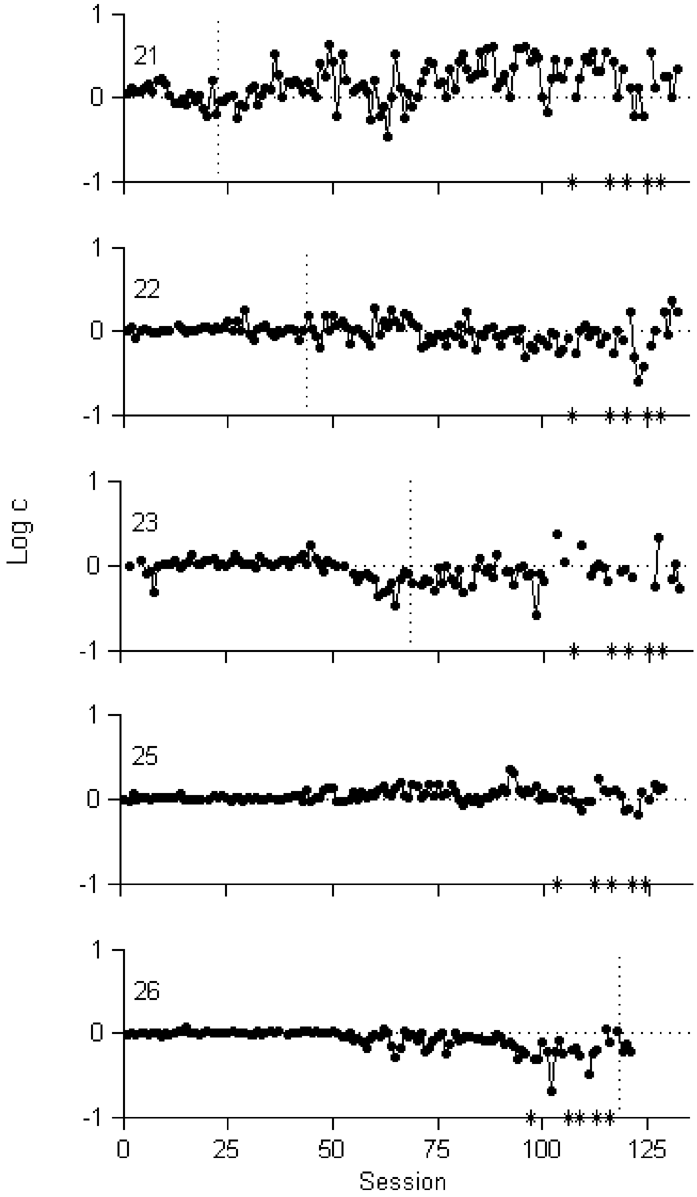


Figure 4.3b. Log *c* estimates over training sessions for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

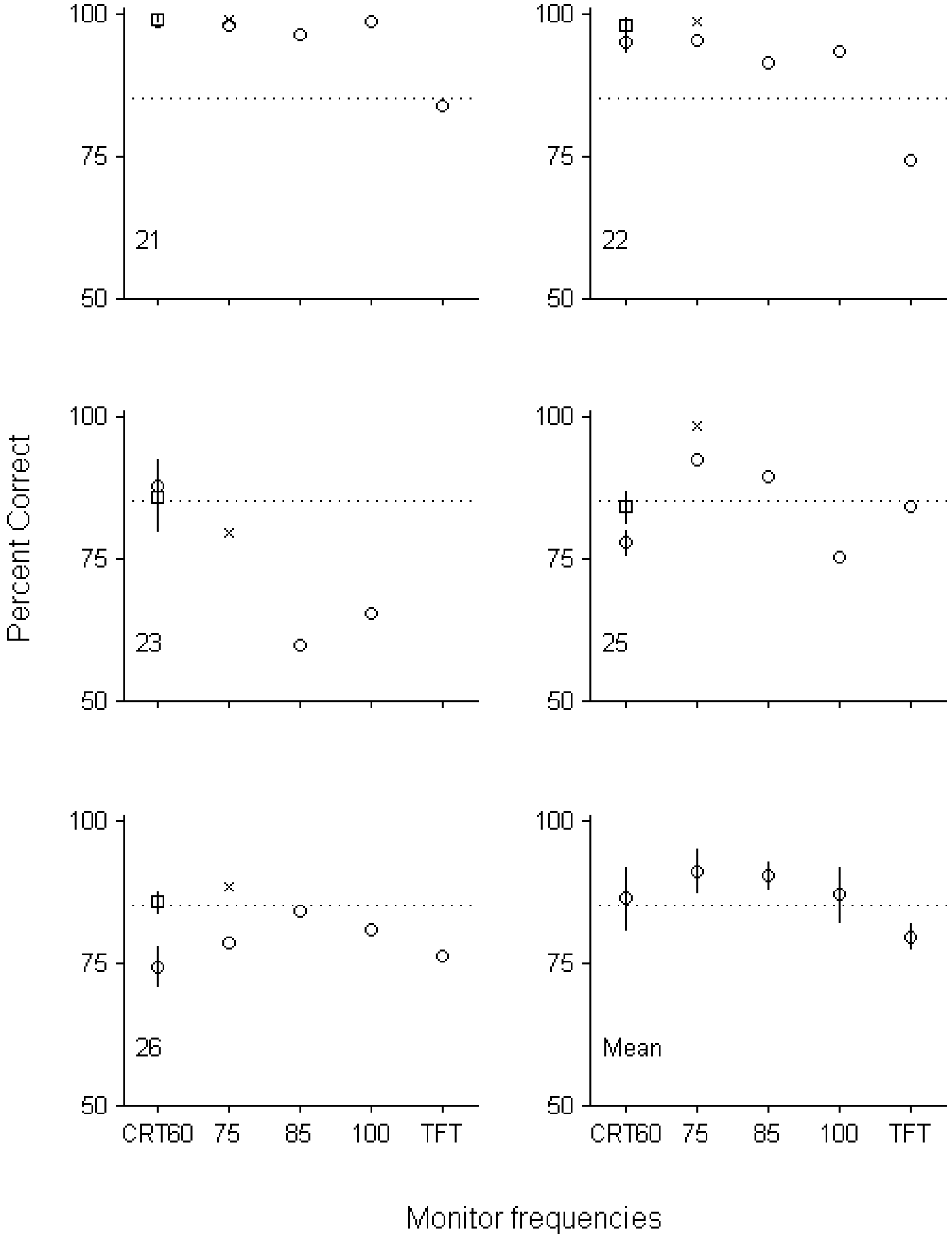


Figure 4.4. Mean percentages correct and standard deviations over the last five training sessions before test sessions began (shown as o) with the CRT monitor and the data from the test sessions. Also plotted are the last five training sessions of the experiment (shown as □). The vertical lines marks one standard deviation each side of the mean and the horizontal dotted line marks 85 %. The bottom right panel shows the average percentage correct and standard error of the mean at each frequency level across those four hens completing all test sessions.

from the test sessions. The second data point is presented for the training data as Hens 25 and 26's accuracy increased after the test sessions began. The data from the replication of the 75 Hz condition are plotted as an x on the graph. The bottom right panel in this figure shows the average percentage correct at each frequency level across those four hens that completed all test sessions (i.e., Hen 23's data were excluded from the averages) and the SEM of these. Hen 21 and 22's accuracy was high during the training sessions before test sessions began with low variability, and their accuracy was high for all the CRT test sessions (range across sessions, 91 to 98 %). However, their accuracy dropped to 84 and 74 %, respectively, when the same stimuli were presented on a TFT monitor. Hens 25 and 26's average percentage correct was 77.8 and 73.8 respectively, over five training sessions before testing but these increased during the testing period to 83.8 and 85.2, respectively (shown as squares). Hen 23 did not respond during two of the test sessions (first 75 Hz session and the TFT session). However, her percentage correct during the training sessions was high (average of 87.5 %) with little variability. Her accuracy was low in the 85 and 100 Hz test sessions (60 and 65 %, respectively), and was only slightly higher than this in the last 75 Hz test session (79 %).

There was no statistically significant difference between the data from the two 75 Hz test sessions (paired samples *t*-test, $t(3)=-2.594$, $p>.05$, $d=-3.588$ (large effect size)). Hence, the averages of these were used in the following analyses. A repeated-measures ANOVA over the data from the five different test sessions was statistically significant ($F(4,12)=3.304$, $p<.05$, $\eta^2=0.524$ (moderate effect size)). However, there were no significant trends for any hen (Theil test, C = 6, 6, 4, 2 and 0 for Hens 21 to 26 respectively, $n=5$, $p>.05$). The average percentage correct with the CRT monitor at the end of training (60 Hz) was not significantly different from the data with the monitor set at 75, 85 or 100 Hz (paired samples *t*-test, $t(4)=-0.896$, $p>.05$, $d=-.451$ (minimum practical effect); $t(4)=0.263$, $p>.05$, $d=.117$ (moderate effect size); $t(4)=0.832$, $p>.05$, $d=.385$ (no effect) respectively). However, there was a significant difference in percentage correct with the TFT monitor and percentage correct with the CRT monitor at 75 Hz and at 85 Hz ($t(3)=4.249$, $p<.05$, $d=-1.544$ (moderate effect size); $t(3)=3.880$, $p<.05$, $d=.516$ (minimum practical effect) respectively). There was no statistically significant difference between percentage correct with the CRT monitor set at 60 Hz, or at 100 Hz and with the TFT monitor ($t(3)=1.035$, $p>.05$, $d=.648$ (minimum practical effect size); $t(3)=-1.176$, $p>.05$, *d*=.210 (no effect)). There was also no significant difference in percentage correct when the CRT monitor was set at 100 and 75 Hz ($t(4)$=2.116, p>.05, d=-1.179 (moderate effect size)), 100 and 85 ($t(4)$=0.0.690, p>.05, d=.323 (no effect)), and 85 and 75 Hz ($t(4)$=1.624, p>.05, d=1.158 (moderate effect size)).

Percentages correct for blocks of 10 trials were examined for each of the test sessions to determine if this changed over a test session. There were 23 test sessions across all hens for this condition, and there were no visual trends in accuracy in the data from 22 of the test sessions. Hen 25 showed a small trend towards increasing percentage correct during the test session with the TFT monitor.

Discussion

All hens' performances transferred from the CRT monitor at 60 Hz to the same monitor set at higher refresh rates. This finding implies that the images presented at the different refresh rates appeared similar to the hens. Of the three possible outcomes considered in the introduction to this experiment, it seems that the third possibility is the best description of these results. That is, the hens learned the discrimination with the flickering stimuli, and also showed transfer to the stimuli presented at higher refresh rates on the CRT monitor than that used in training. It seems that once the hens had been trained to discriminate flickering stimuli at a low flicker rate, they then responded to the steady stimuli as equivalent to the flickering images. Percentage correct with the TFT monitor was lower than with the CRT monitor set at 75 and 85 Hz, but were equivalent to the other refresh rates (60 and 100 Hz).

If discriminating flickering stimuli is more difficult than discriminating steady images, it would have been expected that the hens may have required more sessions to learn the discrimination with the CRT monitor set at 60 Hz, than with the TFT monitor in Experiment 3. The numbers of sessions it took a hen to complete 5 sessions at 85 % or above correct are presented in Table 4.1. Three of the four hens who achieved this took longer with the 60 Hz refresh rate. However, there the defence was not statistically significant (paired-samples *t*-test, $t(3)$ = 1.238, p >.05, d=-.629 (minimum practical effect)). It must be noted however that the sample size is small (n=4), and the variance in the data for Experiment 3 is large, leading to a decreased likelihood of a significant result. Thus, the effect size could give a better

Table 4.1.

The session number at which each hen had responded at or above 85 % for at least 5 sessions.

Hen number	Experiment 2 (TFT)	Experiment 3 (60 Hz)
21	44	22
22	31	43
23	31	68
25	-	-
26	25	118

indication of differences across the groups, however, there was only a minimum effect size showing only a small difference between the groups.

The mean percentages correct and the SEM from the training and test sessions from the condition with the line drawing in Experiment 3 and the present experiment are presented in Figure 4.5 for comparison, plotted against the screen type or refresh rate. For the CRT screen, the average data from Experiment 3 show decreasing percentage correct with decreases in refresh rate, while those for the present experiment appear virtually flat. All averages from the CRT screen sessions in the present experiment are above 85 % correct but the TFT test session gave a lower average. In contrast, only the TFT training sessions average is above 85 % in Experiment 3, with all the CFT test sessions averages lower than 85 % correct. Therefore, it appears that the discrimination was disrupted when transferring from the steady to the flickering stimuli, and was not as disrupted (and sometimes improved) when transferring from the flickering to the steady stimuli. This shows that the transitions between the stimuli are not equivalent and are dependent on the direction of the transfer tested. That is, transfer from steady to flickering stimuli is not the same as transfer from flickering to steady stimuli. This finding also shows that, although hens do not transfer a discrimination from a steady image to the flickering stimuli, they can be trained to discriminate between flickering stimuli, and that this discrimination then transfers to steady images.

In Experiments 2, 3 and 4, all hens showed some level of transfer from the TFT monitor to the CRT monitor and vice versa, indicating that the stimuli appear similar enough to allow transfer to occur. However, the data show that transfer from a steady to flickering stimulus does not occur as readily as transfer from a flickering to steady stimulus.

Accuracy with the TFT monitor was higher in Experiment 3 (the monitor used during training) than in the present experiment. It is possible that accuracy with the TFT monitor would increase with further sessions as in the present experiment the monitor was used to present images for one test session only.

If images are to be used to as substitutions for real objects, there needs to be research into whether hens respond in a similar manner to both images and to real objects. The following series of experiments assesses whether hens transfer a discrimination from objects to photographs of those objects, and vice versa.

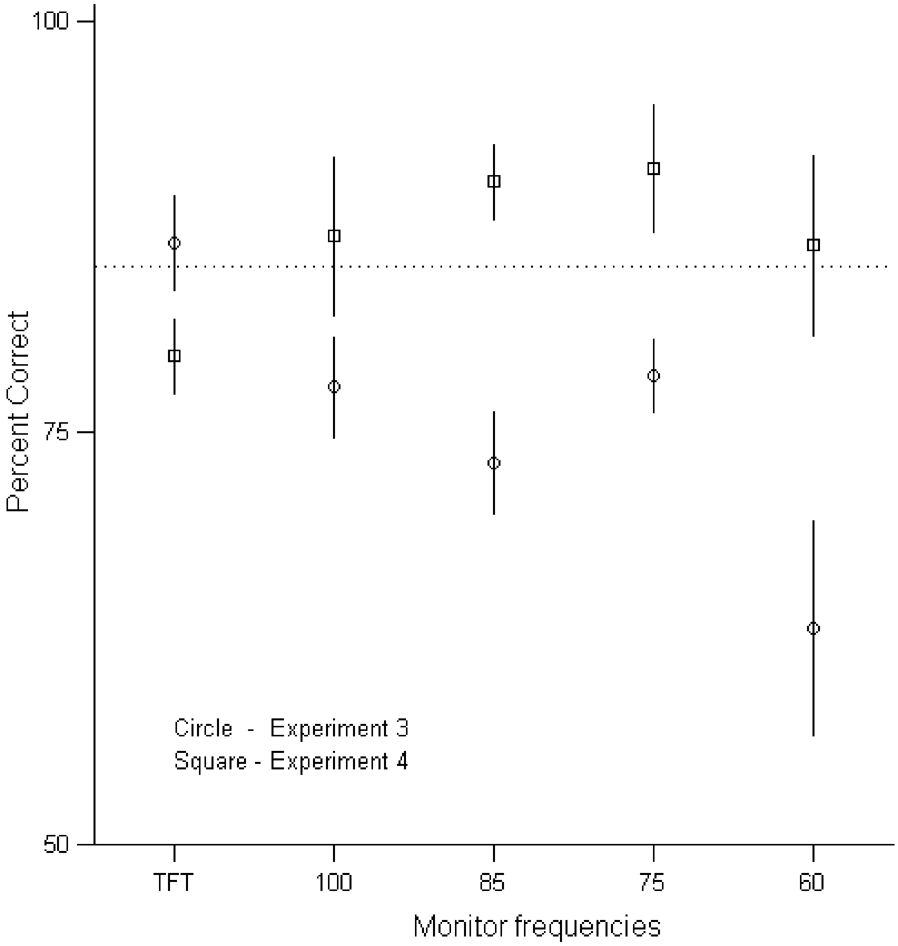


Figure 4.5. Percentage correct and standard error of the means across hens for the line drawing discrimination in Experiments 3 and 4. The dashed horizontal line marks 85 %.

EXPERIMENT 5

As already mentioned, in behavioural research, 2D stimuli are often used as a substitute for 3D objects or animals. However, it is unclear whether animals see pictures or other 2D stimuli as a substitution for the 3D objects being depicted. In order to determine if pictures can substitute objects for animals, research is needed into whether animals respond in a similar manner to both.

Premack (1976) proposed that animals must respond to images in two almost opposing ways if images are to be said to represent objects for animals. Firstly, the animal must respond to the image in the same way that it responds to the object (i.e., it must confuse the object with the image). This provides assurance that the animal's responses to the real object generalise to the image. Secondly, the animal must also respond to the image differently from the way it responds to the object (i.e., the image must represent the object but clearly is not the object). This provides assurance that the animal can discriminate between the image and the object. If it cannot, then the image does not 'represent' the object, but effectively is the object for that animal. However, images are typically used in animal research, not to be a representation of a stimulus in the way Premack suggests, but as a substitute for that stimulus. If animals do see images as a representation, then they will respond to them differently from the way they would to the real stimulus and so the image would not produce the same behaviour as the real stimulus. Thus, when an image is to be used as a substitute for the real stimulus, the animal must respond as if the images and real stimuli are equivalent (i.e., confuse the stimuli).

Watanabe (2000) expanded on Premack's suggestions and proposed four different ways that an animal can show equivalence between 2D stimuli and 3D objects. The first, which he terms correlational, requires that behaviour with a picture is similar to behaviour with the real object (this would seem to be evidence of Premack's first point outlined above). The second, termed interaction, requires that experience with either an object or picture will affect behaviour towards the other. The third, termed distortion, requires that unnatural distortion of a picture will disturb an animal's natural behaviour towards that picture. The fourth, termed direct transfer, is where discrimination of the real objects transfers to the images. The degree of transfer shows a degree of equivalence between the real world and the pictures. Transfer tests were used in Experiments 2-4 to assess if hens showed transfer from objects to photographs, and vice versa. For 2D -3D transfer it is normally the objects and images of the objects that are the focus. Transfer is commonly assessed by training an animal to respond differently to two (or more) objects (or their images) and testing for transfer of this discrimination to images of the objects (or to the real objects). If accuracy remains high during the transfer tests, this shows generalisation across the objects and images, and that the animal can respond to both in a similar manner. However, if accuracy decreases during transfer tests, then the test stimuli are being treated as if they are different from the training stimuli.

While there have been studies assessing transfer of discrimination from objects to pictures across a range of species, the literature with birds is the most relevant to this thesis as it concerns hens' behaviour. Therefore, the research with birds is presented here. There are a few studies with birds that report successful transfer from objects to pictures, or vice versa using transfer tests and several different object types. Cabe (1976) trained pigeons to discriminate between two three-dimensional objects (a white cross and a white rectangular shape) and found that discrimination transferred to black and white photographs and silhouettes, but not to line drawings of those objects. It must be noted, however, that transfer in the opposite direction was not tested. Lumsden (1977) trained a single pigeon using a go/no-go discrimination with differently shaped objects (a wedge shaped and a flat hourglass shaped object) presented from different viewpoints. He tested for transfer to photographs and line drawings and found similar decrements in accuracy across viewpoints were obtained with the photographs as with the objects, but that the level of responding was less for the line drawings. Delius (1992) trained pigeons to discriminate spherical objects from non-spherical objects, and found performance transferred to novel objects and to both photographs and drawings of the objects to varying degrees. The pigeons performed worst with the colour photographs, leading the author to suggest that the photographs may not have accurately depicted colours for the birds and this may have lead to the decrement in performance. In these three cases the pairs of stimuli differed from each other in a number of ways. Such stimuli can be discriminated from each other on the basis of a range of different features, for example, the position of certain points on a stimulus. They raise the question of what aspects of the stimuli were the basis of the original discrimination.

Spetch and Friedman (2006) trained pigeons to discriminate between either two objects or between two images of these objects (examples are shown in Figure 5.1) shown on a TFT monitor. The authors varied the viewpoint from which the stimuli were presented, thus ensuring that the pigeons could not discriminate purely on a single 'cue' of the shape of the object, but were required to learn the discrimination on the shape as a whole. Transfer and reestablishment of discriminative performance was then tested by replacing pictures with objects and objects with pictures for each group of pigeons. The discrimination was also reversed for half of the pigeons so the stimulus associated with reinforcement was now the non-reinforced stimulus, and vice versa. There was some transfer and relatively fast reestablishment of discriminative performance for those pigeons presented with the same reinforced stimulus during training and transfer tests. Those pigeons tested, during the transfer tests, with the opposite reinforced stimulus to that used in training showed more initial disruption and, although performance improved, it remained lower than that of the previously mentioned group. The authors argue that these findings provided evidence that the birds "recognised the correspondence" (p.969) between the objects and the digitized images of them.

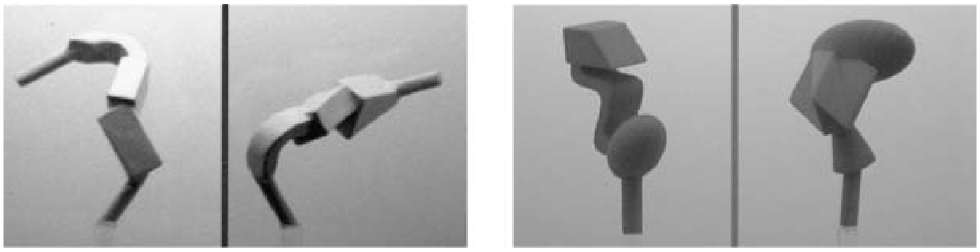


Figure 5.1. Images of the objects used by Spetch and Friedman, (2006) p. 967.

There are a number of other studies that have reported successful transfer between 2D images and 3D objects, however, in these studies, there are factors other than the objects themselves may have contributed to the birds' ability to transfer (such as size, colour and texture). For example, Looney and Cohen (1974) found that pigeons will attack a stuffed or live pigeon when induced with long periods between reinforcement and will similarly attack a photograph, black and white silhouette, and line drawing of a pigeon. However, this behaviour occurred less with the line drawings than with the silhouettes or photographs. The behaviour also depended partly on whether the pigeons had previous experience with the live or stuffed targets. However, it was noted that the pigeons pecked the upper portion of the pictures when they were both the right way up and inverted, therefore it cannot be assumed that the pigeons are seeing the pictorial images as a substitution for the real pigeon.

In another study, Watanabe (1993) found that pigeons' discriminative performance transferred from objects to colour pictures, and vice versa, only when the object was classified as a natural concept by the author (e.g., biologically relevant items such as corn and other food items). Transfer did not occur when the object was classified as a pseudo-concept (e.g., biologically irrelevant items such as coins and other inedible items) indicating that the slides were not equivalent to the real objects. The author concluded that the pigeons were able to classify the stimuli into groups based on whether they were edible or not. However, as Fagot, Martin-Malivel and Dépy (1999) point out, there were size differences between the stimuli could have been a factor in their findings. The pseudo-concept items tended to be larger than the natural concept items. In a later study, Watanabe (1997) trained one group of pigeons to discriminate between real objects (food grains and non-edible items) and their colour photos. Another group were trained to discriminate whether both food and non-food items were real or photographs. Both groups' accuracies transferred to novel stimuli. However, transfer broke down when both stimuli were painted matte black suggesting that colour or texture cues were required for both the real and photograph discrimination.

Although the studies outlined above found some degree of transfer between objects and photographs, they do not provide evidence that birds see pictures as a substitution for real objects, as the birds were most likely using cues other than the objects (e.g., colour) to transfer their discrimination between pictures and objects.

There are also a number of studies that have shown birds either failed to transfer a discrimination from real stimuli to photographs of those stimuli, or failed to respond to photographs similarly to the way they do to the real stimuli. For example, while Candland (1969) found that hens discriminated between slides of conspecifics and concluded that his findings showed hens were able to recognise the conspecifics shown in the slides, Weavers (2000), in a replication of Candland's experiment, found hens' responses to slide images did not generalise to conspecifics, and thus concluded that slide images were not equivalent to real conspecifics for those hens. He argued that the hens in Candland's study most likely did not ‘recognise’ the slide images as real conspecifics. Similarly to Looney and Cohen’s (1974) study outlined above, Weavers found that the hens in his study were attending to the upper parts of the slides rather than to the images represented in the slides. Watanabe, Lea and Dittrich (1993) found that pigeons trained to discriminate real food from “junk” items (e.g., corn kernels from stones) could do the same discrimination when tested with novel foods/items. However, those pigeons trained on colour slides of the same stimuli could not discriminate between slides of novel foods/items. This finding suggests that the slide discrimination was different from that with the real stimuli, and the slides were probably not seen as equivalent the real objects. Bradshaw and Dawkins (1993) trained hens to discriminate between slides of familiar and unfamiliar hens’ heads. The hens discrimination was then tested with novel views of the same hens (e.g., left view of head, front view, tail or feet). Those hens trained with familiar conspecifics failed to show better discrimination in the test sessions from those trained with unfamiliar conspecifics and the authors concluded that the slides were not equivalent to the real hens in these cases. However, it could be argued that the novel stimuli used would be a difficult discrimination regardless of familiarity with the test hen. In addition, although Trillmich (1976) found that budgerigars were able to discriminate live conspecifics and slides of conspecifics in a T-maze discrimination, only one of the birds showed transfer from slides to live conspecifics, and this bird did not show transfer in the opposite direction (from live conspecifics to slides). Dittrich, Adam, Ünver and Güntürkün (2010) trained pigeons to respond differently to different humans during feeding. The pigeons were unable to transfer their discrimination to photographs of the same humans, however, the photographs were not adjusted to be the same size as the humans, and size differences and differing stimulus presentation distances may have accounted for this failure of the pigeons to generalise their earlier discrimination.

The studies outlined above show that the research on whether birds see pictures as substitutions of real objects is equivocal. Bovet and Vauclair (2000) reviewed picture-object recognition and transfer in birds and they note that there are many inconsistencies in the findings reported. It would appear that some bird species (e.g., pigeons) are able to show some transfer between photographs and the objects that they represent, but that this ability is limited, and may be affected by a number of variables (e.g., method of stimulus presentations, type of stimuli used, other cues available to discriminate stimuli). Pictures, therefore, may be inappropriate as a substitute when precise correspondence is required (Delius, Emmerton, Hörster, Jäger, & Ostheim, 2000). The review by Bovet and Vauclair highlights the need for further research on object-picture recognition in animals.

There has been research with hens that has used images as a substitute for the real animal (e.g., Abeyesinghe, et al., 2009; D'Eath & Keeling, 2003; Evans & Marler, 1991; Hauser & Huber-Eicher, 2004; Keeling & Hurnik, 1993; Lundberg & Keeling, 2003). Yet, it is not clear, when using such stimuli, whether hens recognise the relation between objects and their images. Hence, a study assessing whether hens can transfer a discrimination learned with real objects to photographs of those objects (and vice versa) might help to clarify whether pictorial stimuli can be substituted for the real stimulus in behavioural research and add to the body of research presently available.

As repeated trials are required in transfer studies, removing and replacing stimuli can become problematic with real stimuli. For example, while it is fairly simple to replace pictorial stimuli presented on computer or video screens, this becomes more troublesome when objects and photographs are used. One method could be to use a large horizontal wheel that is segmented into a number of sections. Different stimuli could be made visible to the experimental subjects by rotating the wheel so that only the selected stimulus was visible at any one time. In this way, once an animal had been trained to discriminate between different stimuli, they could be tested with photographs of those stimuli and vice versa.

In the studies outlined above, both biologically relevant (e.g., pictures of conspecifics or food) and biologically irrelevant stimuli (e.g., shapes) were used. As previously stated, the problem with the former types of stimuli is that biologically relevant stimuli often differ across a number of features (size, colour, texture, etc). Therefore, transfer can occur between pictures and objects on a number of features rather than a consideration of the stimulus as a whole. To avoid this possibility and ensure that any other cues were ruled out, the stimuli used in the present experiment were children's playing blocks that were similar in colour and texture and were as similar in size as possible, but differed in shape. The objects were lit during experimental sessions by four lights placed in such a way that shadows were eliminated, and any shadows were similarly eliminated in the photographs. This was done to ensure that the photographs and objects appeared as similar as possible.

The present study aimed to establish a procedure that could examine whether hens transferred their discrimination of 3D objects to 2D photographs, and vice versa. It must be noted that the first experiments in this thesis were conducted concurrently with these later experiments, and different stimulus presentation methods were used (i.e., photographs were presented in all following experiments). To orient the hens towards the discriminative stimuli, an observing response was included in the procedure. The hens were trained to break an infra red beam in front of the stimuli continuously for at least 0.5 s before the keys were lit and a choice response could be made. They were also trained to discriminate between two stimuli (green Lego™ playing blocks) or images of these stimuli. A hen was deemed to have learned the discrimination if their accuracy was at or above 85 % correct over five, not necessarily consecutive, sessions. Testing sessions commenced once at least two of the six hens learned the discrimination to 85 (or above) % correct. If all the hens were responding below 85 % correct and the data were not trending after at least 30 sessions, a condition was terminated and the next condition or experiment began.

Pre-testing of the equipment. Prior to the start of this experiment, a group of hens, not involved in the rest of this thesis, went through a series of training and test conditions in an early version of the test equipment. These conditions were intended to train a simple conditional discrimination with one stimulus type (photograph or real object) and test for transfer to the alternative stimulus.

Six hens were split into two groups. One group underwent training with two grey triangular shapes (40-mm long x 32-mm wide x 30-mm high) and two grey rectangular shapes (40-mm long x 27-mm wide x 27-mm high) in the wedge sections of the wheel equipment described below. They were then tested for transfer to photographs. Only one out of the three hens in this group learned the original discrimination to at least 85 % correct, but her performance did not transfer to the photographs.

The other group of hens underwent training with photographs of the grey shapes described above, and were tested with the real objects. All three hens learned the task to at least 85 %, correct, however, none of their performances transferred to the real objects. After the failure in the transfer tests, a series of tests were run to try to establish the source of the existing stimulus control for these hens. These included test sessions with the stimuli presented in both the wedge sections of the wheel that were used in training with which hens had previous experience, and those wedge sections that had not been used in training with which the hens had limited experience.

The percentage corrects for each of these tests are presented in Table 5.1. All of these hens retained some degree of discrimination when the stimuli were removed from the training wedge indicating that some feature of the experimental equipment itself was aiding in the hens' discrimination. In addition, when the hens were tested with the stimuli in the non-training wedge, they showed some discrimination when presented with the original photographs used in training which had some incidental markings – most likely from having been pecked. However, the hens showed no discrimination with unmarked photographs. Thus, the marks on the photographs, rather than the object shown in the photograph, had also been controlling behaviour to some degree. Clearly, data from these hens did not bear on the original question as they never learned the discrimination on the basis intended (i.e., the grey shapes). Therefore, no conclusions about transfer of discrimination were possible as discrimination of the stimuli never occurred. In light of these findings, the equipment was modified for the following experiment by increasing the distance at which the stimuli were presented so that the hens could no longer peck the photographs and objects.

Method

Subjects

Six flock-reared Brown Shaver-Starcross hens (numbered 51-56) served as subjects. At the beginning of the experiment, the hens were one year old, and had some previous experience on simple schedules of reinforcement. The hens were housed individually in wire cages measuring 450-mm long × 450-mm wide × 300-mm high. The room was ventilated and lit on a 12:12-h light:dark cycle with two 100-W incandescent lights. Grit and vitamins were supplied weekly and water was available ad lib within their cages. Throughout the experiment, all hens had red fleshy combs suggesting good health. Each hen was weighed daily before each experimental session (approximately six days per week) and was provided with supplementary feed (commercial laying pellets) if required to maintain them at approximately 80 % (+/–5 %) of their free-feeding body weights.

Table 5.1

Percentage correct for each of the test sessions for those hens that had been trained to discriminate between the photographs.

Test	Condition	Hen Number		
		54	55	56
-	Average last 5 sessions training	90	96	92.8
1	Objects in non-training wedge	38	50	50
2	Objects in non-training wedge	54	43	55
3	Objects in non-training wedge	59	68	48
4	Objects in non-training wedge	38	50	48
5	No photograph in training wedge	79	74	76
6	No photograph in non-training wedge	49	45	49
7	Marked photograph in non-training wedge	62	75	87
8	Unmarked photograph in non-training wedge	55	54	46

Apparatus

A computer (DECpc LPx 433dx) controlled all experimental events and recorded all data using Med-PC® software (Version 1). Total session data were also recorded manually into a data book at the end of each session. The experimental. equipment consisted of two chambers (a stimulus chamber (A) and an experimental chamber (H)) connected by a viewing tunnel as shown in Figure 5.2

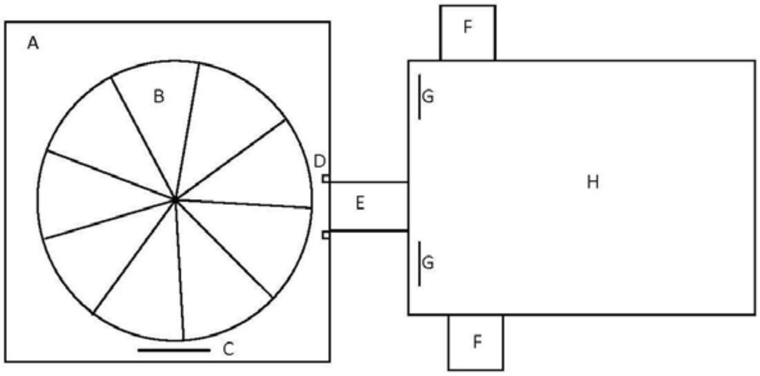


Figure 5.2. A schematic drawing of the experimental equipment. A = Stimulus chamber, B = Wheel, C = Sensor panel, D = Lights in stimulus chamber, E = Viewing tunnel, F = Magazines, G = Response keys, H = Experimental chamber.

The experimental chamber (H) measured 640-mm long × 450-mm wide × 580-mm high and was made of 20-mm thick particle board. The floor was covered with artificial grass matting that could be removed when required and the inside walls of the chamber were painted black. Two red 28v back-lit plastic circular response keys (G) (30 mm in diameter) were surrounded by metal plates and located on the front wall 180 mm apart and 370 mm from the bottom of the chamber. The key lights were multi-chip LED Midget Flange lamps. A brief audible feed-back beep sounded when a response (a force greater than 0.1N) was made on a lit key.

Two magazines (F), containing wheat, were attached separately to each side of the experimental chamber. These were accessible through two square apertures measuring 100-mm wide × 100-mm high. The two response keys were turned off and the magazine was illuminated by a 1-w white bulb following a correct response and during the 3-s access to a magazine.

Located between the response keys, 288 mm from the floor of the experimental chamber, was a viewing tunnel (E) measuring 65-mm wide × 85-mm high that led to the stimulus chamber (A) that measured 900-mm long × 70-mm wide × 32-mm high. The viewing tunnel was 195-mm long and contained two infrared sensors that detected when a hen's head was in the viewing tunnel. Inside the stimulus chamber was a solid horizontal wheel (B) that was 625-mm in diameter that was sectioned into nine equal wedges divided by 10-mm thick plywood walls (as shown in Figure 5.3). The walls were 190-mm high and were painted white (Taubmans™ Exterior white gloss acrylic). Each wedge was 215-mm across at the widest point. Within four of the nine wedges was either two green triangular shaped children's Lego™ blocks with one curved edge (termed the triangular shape) measuring 47-mm long × 31-mm wide × 42-mm high; or two green rectangular shaped children's Lego™ blocks that had been clipped one on top of the other (termed the rectangular shape) measuring 31-mm long × 31-mm wide × 42-mm high (see Figure 5.4). These shapes were attached using Bluetack™ so that they were approximately 20-mm from the front edge. Photographs were taken of the triangular and rectangular objects inside the wedges under two spotlights (to simulate the lighting in the stimulus chamber). These photographs were then adjusted (using Paintshop Pro^{TM}, Version 7) so that they appeared, to the human eye, to be as indistinguishable as possible from the real objects when seen through the viewing tunnel. The colours of the photos were adjusted to be the same as the objects when measured with a Minolta CS-100^{TM} chroma meter. The edges of these photographs were folded and taped in place, 20 mm from the front of the wedge, in four of the wedges (as shown in Figure 5.3). That is, the objects and photographs were alternated across eight of the wedges (one wedge didn't contain any stimuli and was not used in the experimental sessions). Figure 5.3 illustrates how the objects and photographs were positioned in the wedges. The wheel could be rotated using a motor and chain which was under control of the computer.

The end of the wall of each wedge section was painted black (Painters touch™, flat black) as can be seen in Figure 5.5, and a small square of tin-foil was pasted at a different height on each wedge. A panel of nine infrared sensors (C), set at corresponding heights to the tinfoil squares, was screwed to the inside of the stimulus chamber. This allowed the computer to locate each individual wedge and to

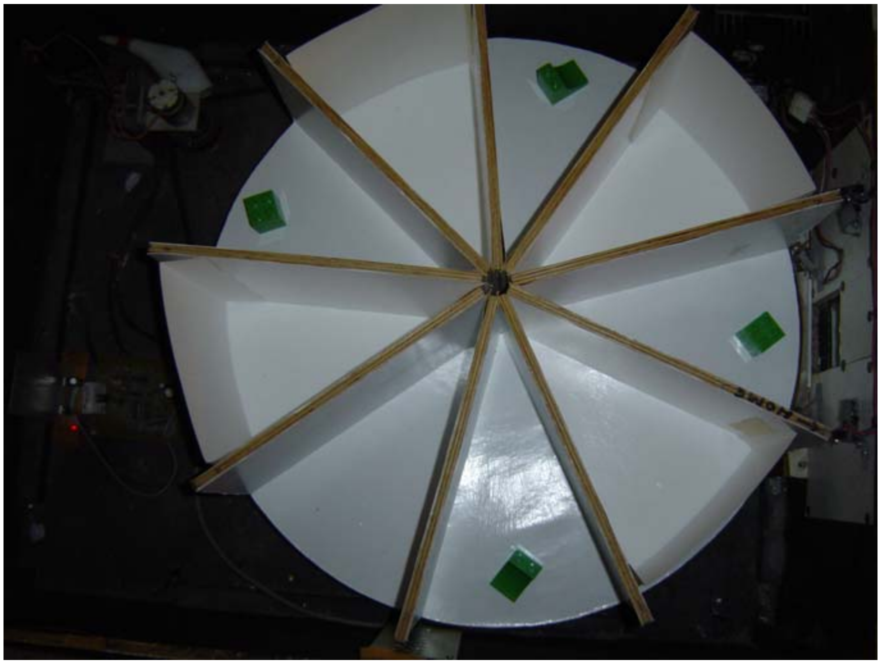


Figure 5.3. Aerial view of the experimental wheel showing the placement of the shapes and photographs.

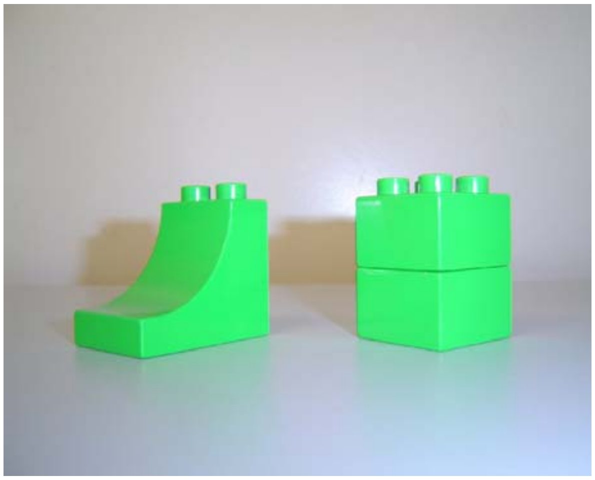


Figure 5.4. Image of the Lego^{TM} shapes used as discriminative stimuli.

Figure 5.5. Side view of the stimulus chamber and the experimental chamber in the background.

stop a selected wedge directly in front of the viewing tunnel where four 5v white halogen bulbs (D) could be lit to illuminate the wedge.

Procedure

Beam-break training. The method of successive approximations was used to shape the behaviour of pecking a lit key in a single key experimental chamber different from the one used in experimental sessions. Once all the hens were pecking the key reliably, they were then trained in the experimental chamber to place their head in the viewing tunnel and break the sensor beam, again using the method of successive approximations. A feedback beep was sounded when the beam had been broken for 0.1 seconds and one magazine was operated. The magazine used (left or right) was alternated from trial to trial. During this time, the key lights were not lit. Once all the hens were reliably and quickly breaking the beam, one key was lit on each trial, and the hens were required to break the beam and then peck the lit key before a reinforcer was available from the magazine associated with that key. Again, the key light and magazine that operated alternated from trial to trial. All responses to the lit key resulted in a reinforcer. The hens received two sessions of 30 reinforcers at this training level. Next, the discrimination training phase began.

Discrimination training. Hens 51-53 were trained with the Lego^{TM} objects as the target stimuli, and Hens 54-56 were trained with the photographs of the Lego^{TM} objects as the target stimuli. During discrimination training, the time required for an effective beam-break was gradually increased to 0.5s.

Figure 5.6 outlines the procedure in a flowchart. Each session began with a 10 s ITI period. This ITI period between trials was necessary to allow time for the wheel to move to the correct position and present the next stimulus to the hen. On a small number of trials, the wheel took longer than 10 s to move to the next stimulus wedge and on those occasions the trial did not begin until the wheel was in place and as a result the ITI was slightly longer than 10 s on these trials. When the wedge was situated in front of the viewing tunnel (and the 10 s ITI period had elapsed), the lights in the stimulus chamber were turned on.

Once a hen had broken the sensor beam in the viewing tunnel for at least the required time and the beam was no longer broken (i.e., the hen had removed her head), the two keys were lit. For the first ten sessions only the correct key was lit. The keys remained lit until either a response had been made, or the trial was aborted. The hen was required to peck the response key within 5 s of removing her head from the viewing tunnel to ensure all choices were made shortly after viewing. If the hen failed to peck the key (or re-break the beam) within this time, the trial was aborted, the lights in the stimulus chamber and the key lights were turned off and a new trial started. Re-breaking the beam within the 5 s period initiated a further 5 s for a key peck to occur.

Following a correct response (right key response following a triangular shape presentation, left key response following a rectangular shape presentation), the magazine and magazine light was operated for 3 s if a reinforcer was due. If no reinforcer was due, only the magazine light was operated for 3 s. Incorrect responses resulted in a 3 s blackout period before the next trial began.

Whether the triangular or rectangular stimulus was presented was controlled pseudo-randomly according to a predetermined series that ensured that no more than three triangular or rectangular trials occurred consecutively and that the total number of each trial type would be approximately equal within a session. The continuous reinforcement used in initial training was replaced by a variable ratio (VR) schedule of reinforcement so that reinforcers were scheduled for a correct response after a variable number of correct responses. When a reinforcer was not scheduled for that

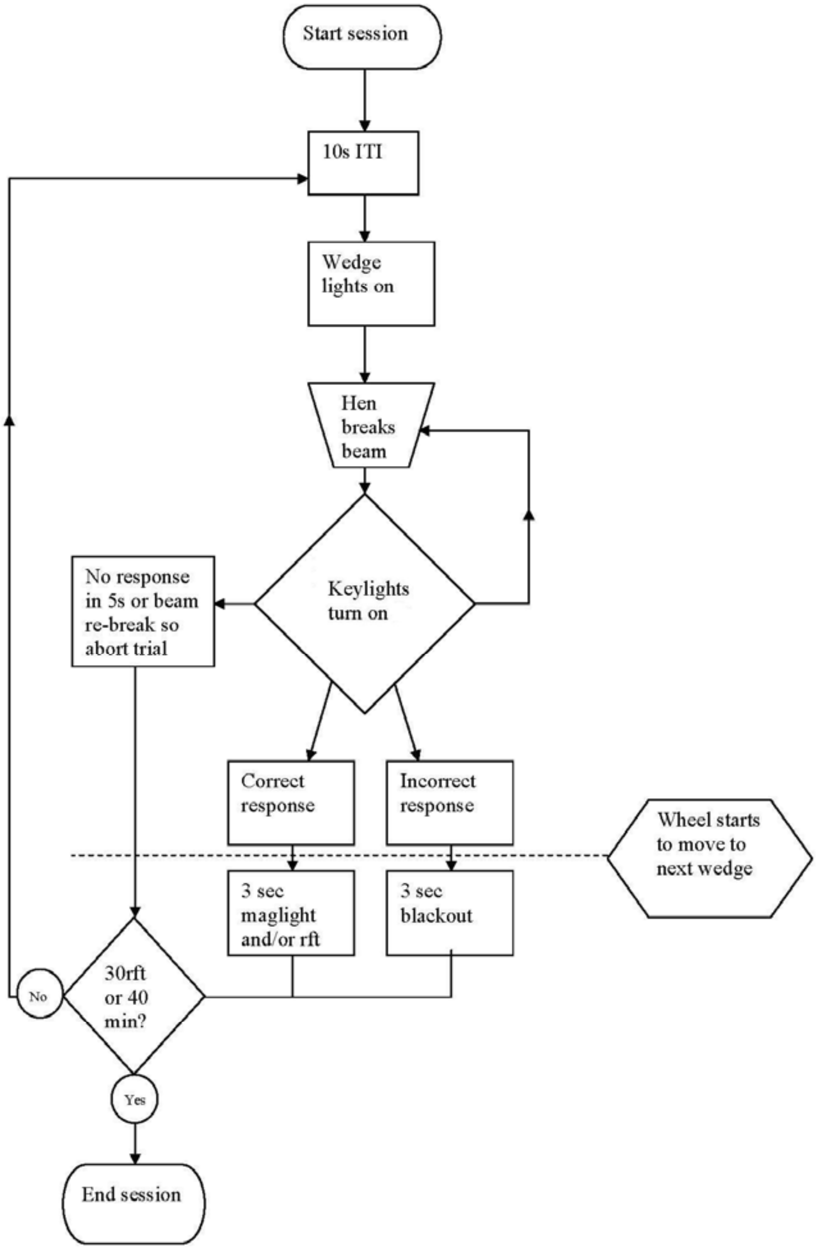


Figure 5.6. Flowchart diagram of the procedure.

trial, a correct response resulted in 3-s illumination of the magazine light, but the magazine was not raised. This VR requirement was gradually increased over a period of 29 sessions until all hens were responding under a VR 2 schedule of reinforcement. Decisions to increase the VR were based on the combination of relatively stable performance and the subject having received approximately 30 reinforcers during the experimental session.

In addition to the VR 2 requirement, reinforcer delivery was controlled so that the rate of reinforcement was equal across the two stimuli. Reinforcers were allocated to a trial type pseudo-randomly. For example, if the next reinforcer was scheduled for a correct response on a triangular trial, any correct responses on rectangular trials were followed by the magazine light, but no food presentation, until the scheduled reinforcer for a correct response on the triangular trial had been collected. Correct responses of both types counted towards the VR requirement of the pre-selected trial type (triangular or rectangular shape) and the reinforcer was delivered for a correct response that completed that VR. By controlling the rate of reinforcement in this manner, the VR requirement was sometimes slightly increased, particularly if the hen responded mainly to one key.

Sessions were terminated after 40 min had elapsed or after 30 reinforcers had been obtained, whichever occurred first.

Results

Occasionally the hens laid eggs during experimental sessions resulting in atypical performance and therefore data from these sessions were discarded. In addition, data were discarded if there were equipment problems during the session (such as a blown light bulb or broken keys).

Hens 51-53 were presented with the objects and Hens 54-56 were presented with the photographs of the objects. The left panel of Figures 5.7a and 5.7b presents the overall percentage correct for the training sessions. After extensive training, none of the hens' accuracies were above 85 % and there were no upwards trends. After 62 sessions of training, most of the hens' performances were at chance levels, except for Hens 51 and 54 who showed some discrimination. There were a minimum number of 100 trials in a session, therefore any percentage above 59 % was significantly different from chance at a .05 level of significance (binomial test). Over

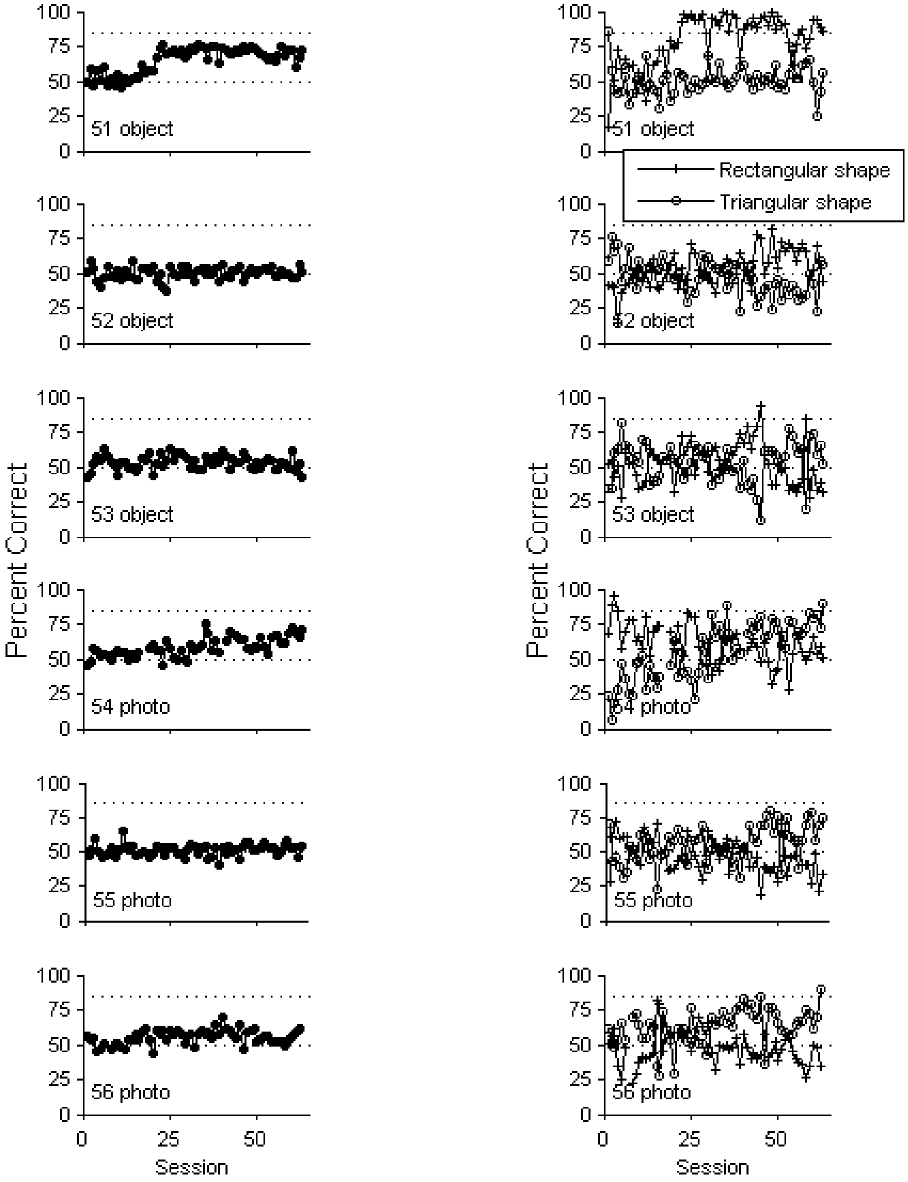


Figure 5.7a. The left panel shows the overall percentage correct across training sessions for all six hens. The right panel shows the percent correct to each of the stimuli. The horizontal dotted lines mark 50 % and 85 % correct.

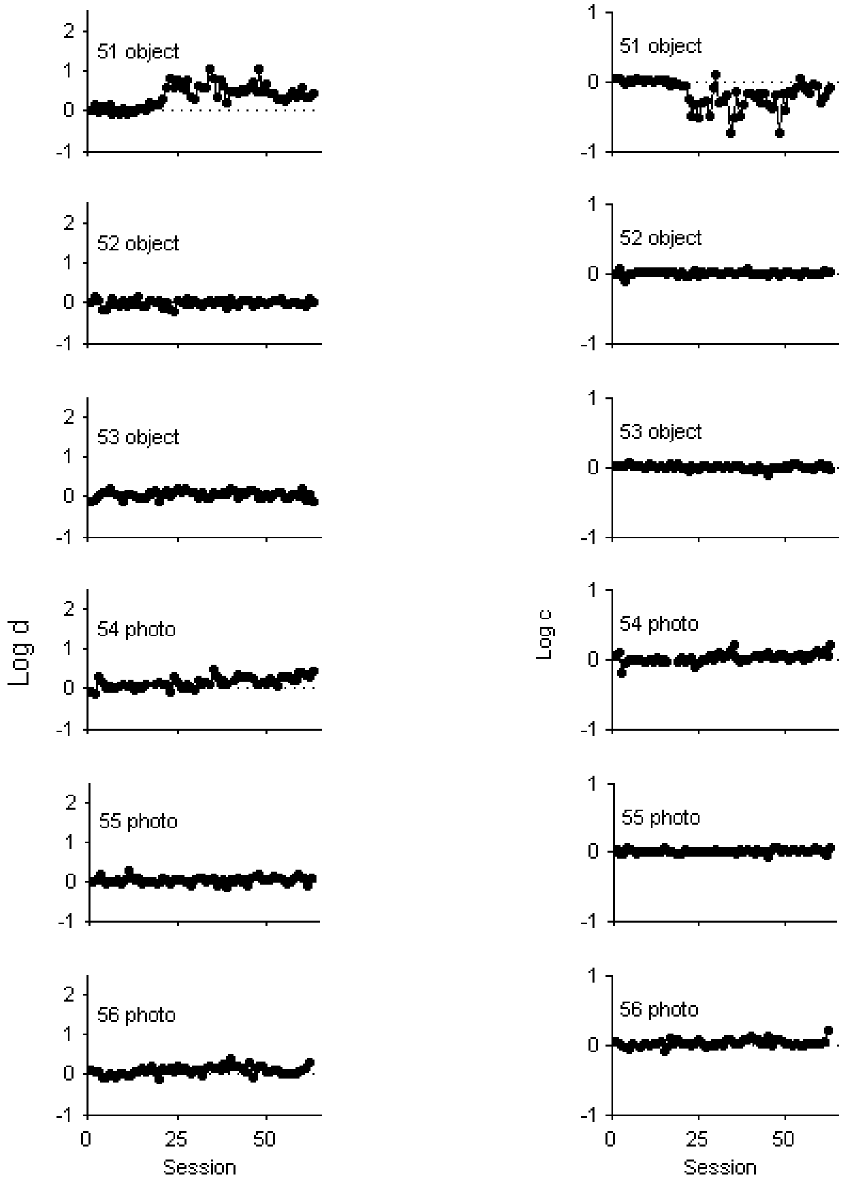


Figure 5.7b. The left panel shows the log *d* estimates across training sessions for all six hens. The right panel shows the log *c* estimates across sessions.

the last five training sessions, Hens 51 and 54's percentages correct ranged between 60 and 72 %, significantly above chance, but they were not trending upwards. On average, over the last five training sessions, the hens broke the beam 1.4 to 2.8 times across trials. That is, they generally did not break the beam for the 0.5 s required to complete the observing response on the first attempt. Over the last five training sessions, the median amount of time from the first beam break to a key peck ranged from 1.9 to 3.8 s over all hens.

The right panel of Figures 5.7a and 5.7b presents the percentage correct on trials with each stimulus, and the log *c* estimates across sessions. Hen 51's accuracy with the rectangular shape ranged from 81 to 94 % correct over last five training sessions, significantly different from chance using the binomial test (above 62 % using a minimum of 50 trials). However, Hen 51's percentage correct remained around chance (50 %) with the triangular shape. The log *c* estimates show she had a left key bias. None of the other hens showed any biases towards responding more on one key than the other. Hen 52 showed slightly higher accuracy with the rectangular shape (range, 45 to 70 % over the last five training sessions) than with the triangular shape (range, 23 to 59 % over the last five training sessions). In contrast, Hen 53 showed higher accuracy with the triangular shape (28 to 49 % over the last five training sessions) than with the rectangular shape (52 to 74 % over the last five training sessions). Hens 54-56 were trained with the photographs of the objects. These three hens all showed higher accuracy with the triangular shape (62 to 90 % over the last five training sessions) than with the rectangular shape (range 21 to 65 % over the last five training sessions) demonstrating a bias towards selecting the right key

Discussion

The aim of this experiment was to establish a procedure that could examine whether hens showed transfer of a discrimination from 3D objects to 2D photographs, and vice versa. However, none of the hens attained high levels of discriminative performance and for most hens, accuracy remained close to chance levels (50 %). As a result of this, transfer to the alternative stimulus could not be assessed. Also, all hens showed strong biases in their responding, but they were in different directions. It seems that these hens had difficulty discriminating between these objects when shape was the only cue to distinguish them.

While Hens 51 and 54 showed some level of discrimination, as their responding was consistently above 50 %, these hens also showed response biases (unequal responding to the two stimuli). As stated previously, DeMello (1989) suggested that performance to each stimulus should be examined separately in conditional discrimination tasks in order to determine the pattern of responding that underlies the overall percentage correct. The hens in the present experiment showed biases to one key, as accuracy with one stimulus was consistently higher than accuracy with the other stimulus. However, these biases were not in the same direction for all hens. These differences displayed by the hens imply that the biases were not a result of consistent extraneous variables, such as differences in the force necessary to trigger a key or level of feed available in a magazine presentation. It is unclear what led to these biases. Although it is tempting to say that only one stimulus gained control of the hens' behavior, this seems unlikely as the biases were different across birds. For example, Hen 51 was mostly correct with the rectangular shape (showing a left key bias) and Hen 54 was mostly correctly with the triangular shape (showing a right key bias).

It is possible that the experimental equipment, stimuli, or procedure (or all three) used were unsuitable for examining discrimination in hens. However, it seems unlikely that the procedure was problematic as it is known that hens are able to do conditional discrimination tasks as shown in the previous experiments reported here and in other published studies (e.g., DeMello, Foster, & Temple, 1992; Temple, Foster & O'Donnell, 1984). Some researchers have stated that successive procedures (as used in the present experiment) are more difficult to learn (McLean & White, 1983), and that performance is more accurate and stable than with simultaneous procedures (Bushnell, 1999; Miloševič, 1993). However, Shelton, Picardi and Green (1982) found little or no effect on performance across these methods of stimulus presentation, and the findings from Experiment 1 of this thesis show that the two presentation methods gave similar results. Therefore, the successive stimulus presentation used in this experiment still seemed warranted.

Another possible reason for the hens' difficulty in learning the discrimination may be the size of the required observing response. In Experiment 3, it was shown that the addition of an FR 5 to the discrimination task improved all hens' accuracies. The observing response in the present experiment consisted of breaking a beam for 0.5 s. It was found that all hens repeatedly broke the beam, and therefore were repeatedly in a position to see the stimuli. The median amount of time that each hen took from first beam break to key peck ranged from 1.9 to 3.8 s. In contrast, the median amount of time taken from first FR response to key peck ranged from 3.7 to 6.9 s, which is almost twice as long as in the present experiment. These differences in times suggest that it is possible that increasing the length of time that the beam needed to be broken may have lead to greater accuracy by these hens.

A further possible reason could be that in the redesign of the early prototype of the experimental equipment (outlined above), the viewing tunnel that the hens needed to place their head in to view the stimuli was extended so that the hens could not peck and mark the objects and photographs. However, it is possible that the stimuli were now be too far away for the hens to view. For example, Dawkins and Woodington (1997) found that hens that had been trained to move toward objects viewed at different distances were more accurate at discriminating the stimuli when they were allowed to view them at close distances (5-25cm) than when they were required to choose at longer distances (120cm). They also found that hens were more accurate at transferring a learned discrimination between differently coloured objects to photographs of those objects when they were viewed at short distances. The authors suggest that important information for hen recognition may be lost at greater distances. Weavers (2000) assessed hens' discrimination of slides of conspecifics at two distances. Initially the slides were presented at a distance of 600mm, but this was later moved to 150mm. It was found that this change in viewing distance initially disrupted the hens discriminative performance for some hens, however, performance was regained after further training.

One way to determine if the experimental equipment or the stimuli led to the failure of the hens to learn the discrimination would be to conduct a further experiment with stimuli that differ on a dimension that hens have been shown to be able to discriminate. One such dimension is colour. In the next experiment the colours of the stimuli were changed, all other factors were kept the same (including the shapes of the stimuli). If the hens failed to learn this colour discrimination, this would imply that some aspect of the experimental equipment itself may have hindered an otherwise simple discrimination. For example, the stimuli may have been presented at a distance that made it difficult for the hens to see, or the observing response may have needed to be increased to ensure the hens were in the presence of the stimuli for longer. Alternatively, if the hens learned this discrimination, this could indicate that the experimental equipment used in the present experiment was appropriate, but that the stimuli themselves were difficult to discriminate.

EXPERIMENT 6

There were four aims of this experiment. The first aim was to assess if hens could learn to discriminate between two stimuli that differed in colour as well as shape. The objects selected were a green and a red coloured Lego^{TM} block of the same shapes as in Experiment 5. As stated previously, hens are able to discriminate easily between colours and many studies have used colour as a discriminative stimulus. For example, Jones, Carmichael and Rayner (2000) found that chicks preferred white or yellow string pecking devices to red, green or blue ones. Dawkins and Woodington (1997) trained chickens to discriminate between blue and red objects. Foster, Temple, Mackenzie, DeMello and Poling (1995) found that hens were easily able to distinguish between red and green keys used as sample and comparison stimuli in a matching-to-sample task. The procedure used in the present experiment was the same as in Experiment 5 but with red and green shapes as stimuli.

If the hens learned this colour/shape discrimination, then the next aim was to examine whether hens transferred that discrimination from 3D objects to their 2D photographs, and vice versa. As stated previously, there have been some studies (Looney & Cohen, 1974; Watanabe, 1993, 1997) that have found pigeons can transfer a discrimination when the stimuli differed across a number of features (e.g., size, texture). Thus, to reduce the number of features that could be used to discriminate the stimuli in the following experiment, the stimuli differed on one or two features only (colour and/or shape).

Given that it was possible that the stimuli that were presented at a distance may have made it difficult for the hens to see them in Experiment 5, the third aim was to establish whether performance was disrupted by moving the stimuli so that they could be viewed at a closer distance by the hens. The fourth and final aim was to establish whether the hens could learn to discriminate between the two shapes of the same colour that they had previously failed to learn in Experiment 5 when the stimuli were moved and could be viewed from a closer distance.

Method

Subjects

The subjects were the same as those used in Experiment 5. The housing conditions were the same as Experiment 5 except that, between Conditions 1 and 2, the individual cages housing the hens were increased to 500-mm long × 510-mm wide × 420-mm high.

Apparatus

The apparatus was the same as that used in Experiment 5. In Conditions 1 and 2, the stimuli were the same as those used in Experiment 5, except that the rectangular shapes were red (see Figure 6.1). In Condition 3, the stimuli were identical to those used in Experiment 5. The photographs of the stimuli were adjusted in the same manner as used in Experiment 5, so that they were the same size and colour as that of the objects. For Conditions 2 and 3, the viewing tunnel was shorted to 65-mm long, and a removable glass panel (100-mm wide × 150-mm high) was fitted at the end of the viewing tunnel to prevent the hens from reaching and marking the stimuli. This glass panel was cleaned at the beginning of each session for each hen.

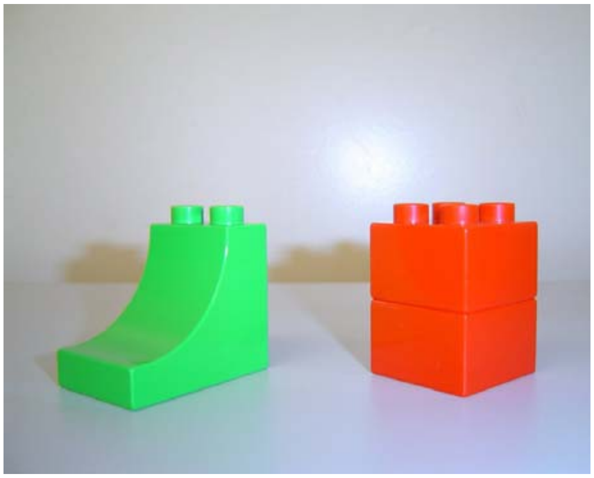


Figure 6.1. Photograph of the green and red Lego™ shapes.

Procedure

Condition 1. Sessions were the same as the discrimination training sessions used in Experiment 5. When a red shape was the target stimulus, a peck to the left response key was correct, and when a green shape was the target stimulus, a peck to the right response key was correct. Hens 51-53 were trained with the objects, and Hens 54-56 were trained with photographs of the objects. The VR requirement was increased over a period of 35 sessions until all hens were responding on a VR3 schedule of reinforcement. Decisions to increase the VR were based on a combination of relatively stable performance and the subject having received approximately 30 reinforcers in less than 40 minutes during the experimental session.

There were a total of four test sessions in Condition 1. Test sessions began after at least two hens were responding at or above 85 % accuracy for five sessions, or were responding above chance. There were a minimum number of 118 trials for these hens over the last five sessions of training before test sessions. Therefore, any percentage above 57 % was significantly different from chance at a .05 level of significance (binomial test). Over the last five sessions of training before the first two test sessions, Hens 52 and 55's performances ranged between 61 and 76 % which was significantly above chance.

Two test sessions occurred after the first 60 training sessions and, due to experimental equipment problems, two test sessions occurred after a further 60 training sessions. Hens received at least two training sessions between every test session.

Condition 2. Sessions were the same as Condition 1 except that the objects were now closer as a result of shortening the viewing tunnel. The beam break requirement was reduced at the beginning of this condition to 0.3 s, and was increased back up to 0.5 s after 14 sessions. The test sessions began after a minimum of 72 training sessions. There were two test sessions for all hens with one training session between test sessions.

Test sessions, Conditions 1 and 2. Test sessions followed the same procedure as used in discrimination training, except that the alternative stimulus was presented to the hens. That is, those hens that were trained using the real objects were presented with the photographs of those objects, and those hens that were trained with the photographs were presented with the real objects.

Condition 3. Sessions followed the same procedure as in Condition 2 except that the objects were the two differently shaped green objects as used in Experiment 5, and the beam break requirement started at 0.5 s. The correct response was to peck the right key following a triangular shape presentation, and to peck the left key following a rectangular shape presentation. Test sessions began after at least two hens were responding at or above 85 % accuracy for five sessions, or were responding above chance.

Test sessions, Condition 3. There were 12 test sessions in this condition. Table 6.1 presents the variables that were changed in each test session, and the order in which the test sessions occurred. Four of these sessions were transfer tests in which the alternative stimulus set was presented to the hens (that is, the photographs were presented to the hens trained with the objects [Hens 51-53], and the objects were presented to the hens trained with the photographs [Hens 54-56]). In two of these, the test stimuli were presented in the wedges that that hens had experienced in training (Type 1 tests), and in the other two, they were presented in the wedges not used for training with that hen (Type 2 tests). The objects were placed in four of the nine available wedges, and only these wedges were used to present the stimuli during discrimination training – these are referred to as the training wedges. A further four wedges were used to present the stimuli to the hens trained with the photographs – these were the training wedges for these hens. One of the wedges was not used during training or testing. In the Type 1 tests the test stimuli were presented in that hens training wedges, and in the Type 2 tests they were presented in the wedges not used with that hen during training.

A further four test sessions were used to test for possible unintended sources of stimulus control, in that the stimulus set that the hens had been trained to discriminate were presented to the hens, for two of the test sessions, in the wedges that were not used during training for that hen (Type 3 tests). In the other two of these test sessions, the training stimuli were presented in the wedges that had been used for training with that hen, however, the edge of the wheel was masked from view (Type 4 tests). Further, as it may have been possible for the hens to discriminate based on differences in the appearance of the particle board the equipment was made of rather than on the basis of the stimuli, the bottom edge of the wheel was either visible or covered by a piece of white Perspex (measuring 10-mm high × 115-mm wide) during some test sessions.

In the final four test sessions, to check for unintended sources of stimulus control, no stimuli were presented to the hens. That is, the objects were removed from the wedges for those hens trained with the objects, and photographs of a white background (i.e., a photograph with the stimuli removed) were presented for those hens trained with the photographs. All four of these test sessions were conducted using the wedges that the hens had used during training. However, in two of these

Table 6.1.

Table showing the variables that were changed in each test session in Condition 3 and the order the test session were presented in. There were six different types of test sessions numbered 1 – 6, and each type occurred twice. In some sessions, the training stimuli (either the objects or photographs) were used, and in other sessions the alternative stimulus set (objects for those trained with photographs and photographs for those trained with objects - the non-training stimuli) were used. In some sessions, no stimuli were presented. In a number of sessions the stimuli were presented in the sections of the wheel used for that stimulus in training (the training wedge) and in others the stimuli were presented in sections not used during the training for that hen (the non-training wedge). In two sessions, the edge of the wheel was masked from view.

Test Session		Stimulus Set	Section of Wheel	Edge Masked
Order	Type			
1	A	Non-training	Training	No
2	C	Training	Non-training	No
3	E	No stimuli	Training	No
4	D	Training	Training	Yes
5	F	No stimuli	Training	Yes
6	B	Non-training	Non-training	No
7	D	Training	Training	Yes
8	F	No stimuli	Training	Yes
9	C	Training	Non-training	No
10	A	Non-training	Training	No
11	B	Non-training	Non-training	No
12	E	No stimuli	Training	No

sessions, the edge of the wheel was visible (Type 5 tests), and in the other two sessions, the edge of the wheel was masked (Type 6 tests).

Results

Hens 51-53 were presented with the objects and Hens 54-56 were presented with the photographs of the objects. Figures 6.2a and 6.2b show the overall percentage correct, and the log *d* estimates, over all training sessions for Condition 1. Hens 51, 53, 54 and 56's percentages correct were all above 85 % for at least five sessions within 52 training sessions, and their performances remained consistently stable and above 85 % correct for all remaining training sessions in this condition. Hens 52 and 55's percentages correct continued to increase after test sessions began. Hen 55's percentage correct was at or above 85 % correct for five sessions after the first set of test sessions. Hen 52's percentage correct eventually stabilised at around 85 % but was not at or above this for five consecutive training sessions before the end of this condition.

Figures 6.3a and 6.3b present the percentage correct for each stimulus, and the log *c* estimates respectively for Condition 1. Hens 52 and 53 show no consistent differences in accuracy of responding across the two object stimuli, however Hen 51 showed slightly more biased responding to the left key (red shape). Hen 54 showed no consistent bias towards responding more on one key. Hens 55 and 56 showed biased responding to the right key (green shape) in the later training sessions.

Figure 6.4 shows the mean percentage correct and standard deviation over the last five training sessions and the data from the four test sessions, for all six hens. Mean 1 shows the average percentage correct of the last five training sessions before the first test session. As there were a large number of training sessions between the second and third test sessions, and as Hen 52 and 55's accuracy had increased between these test sessions, a second mean (Mean 2) was calculated and this shows the average percent correct for the last five training sessions before the third test session. Hens 51, 53, 54 and 56's percentages correct remained high (above 85 %) during the test sessions with only one exception (the 3rd test session for Hen 53 at 75 %). Hen 52's percentages correct in the test sessions ranged from 64 to 73 % and were generally lower than her percentage correct during the training. Hen 55's percentages correct over test sessions ranged from 81 to 92 %, higher than during the training sessions. A repeated-measures ANOVA showed no statistical difference

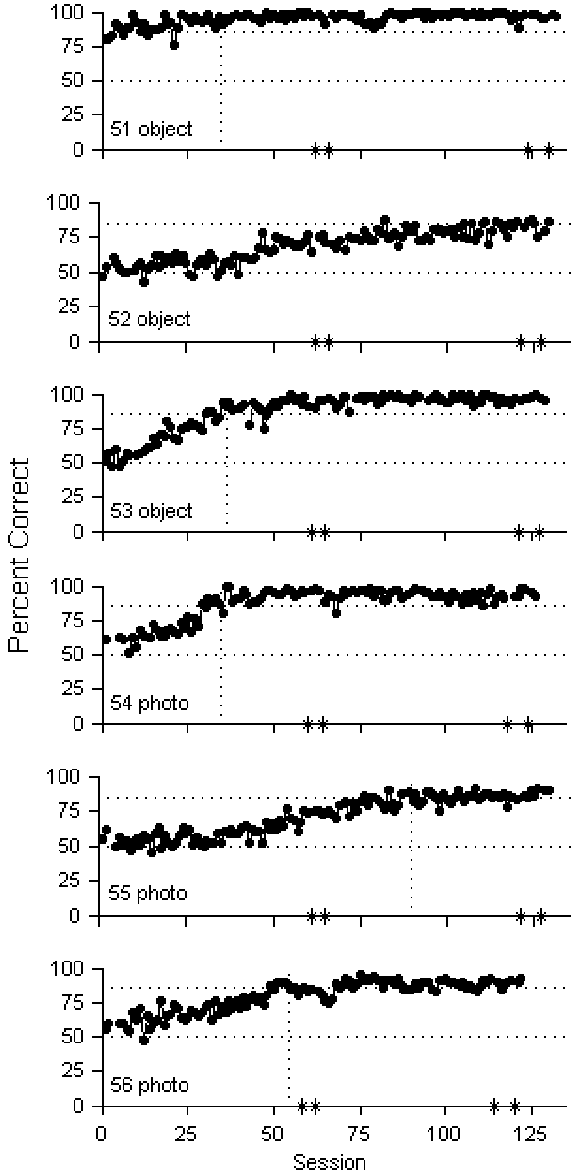


Figure 6.2a. Overall percentage correct over the training sessions with the red and green shapes (Condition 1) for all hens. The vertical dotted line marks the point at which each hen had responded at or above 85 % for five sessions. The horizontal dotted lines mark 50 % and 85 %. The locations of the test sessions are indicated by an asterisk on the x axis.

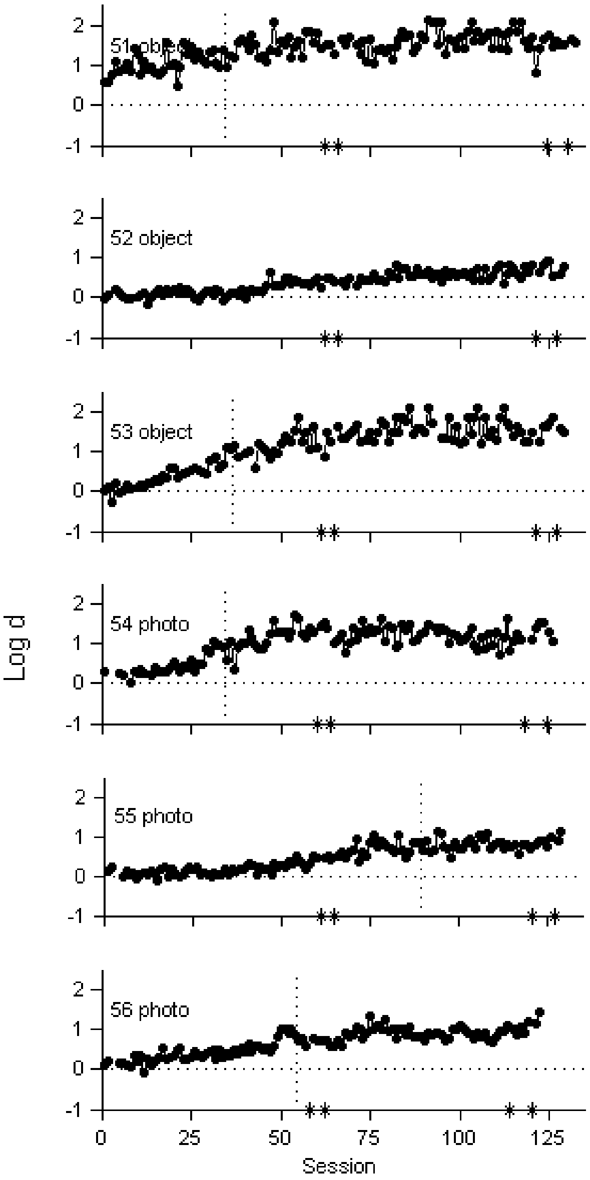


Figure 6.2b. Log *d* estimate plotted against session number for all hens. The vertical dotted line marks the point at which each hen had responded at or above 85 % for five sessions. The locations of the test sessions are indicated by an asterisk on the x axis.

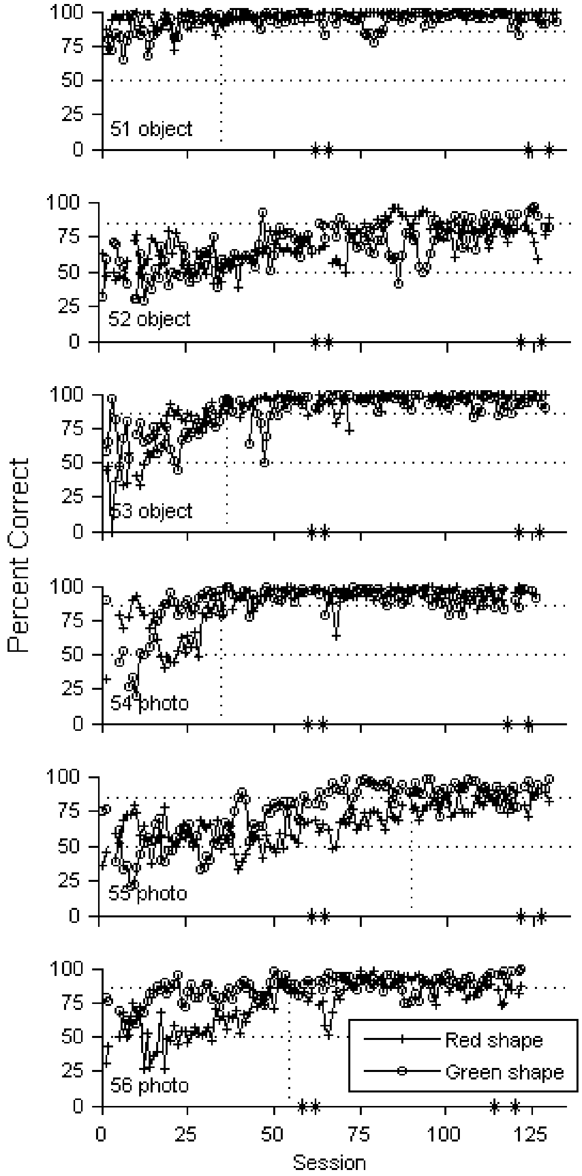


Figure 6.3a. Percentage correct for each stimulus over training sessions for all hens (Condition 1). The vertical dotted line marks the point at which each hen had responded at or above 85 % for five sessions. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the locations of the test sessions.

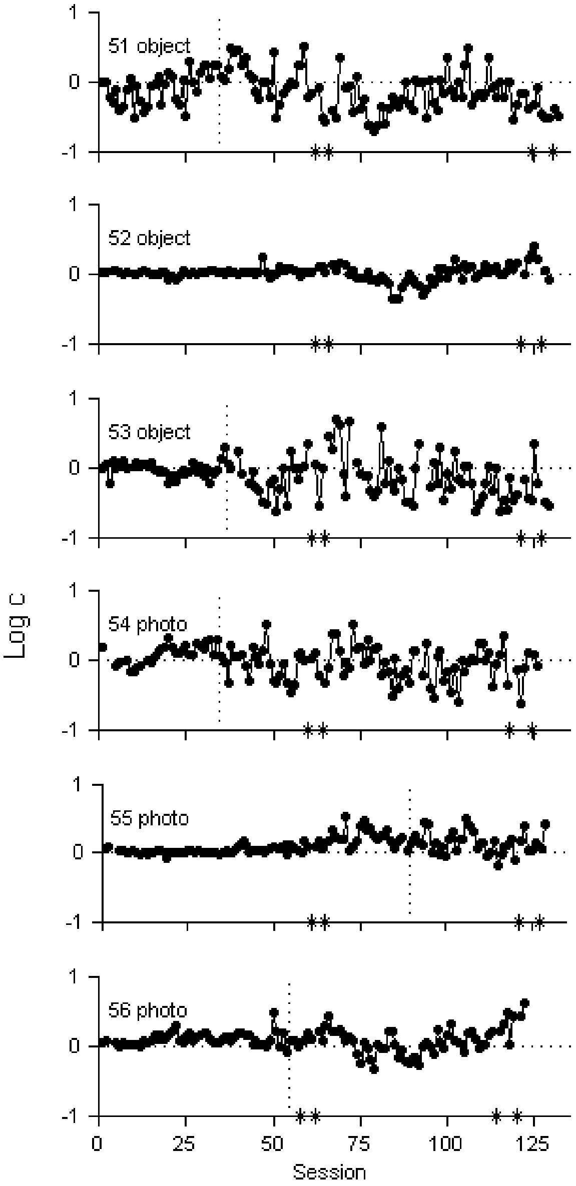


Figure 6.3b. Log *c* estimates over training sessions for all hens (Condition 1). The vertical dotted line marks the point at which each hen had responded at or above 85 % for five sessions. The asterisks on the x axis represent the locations of the test sessions.

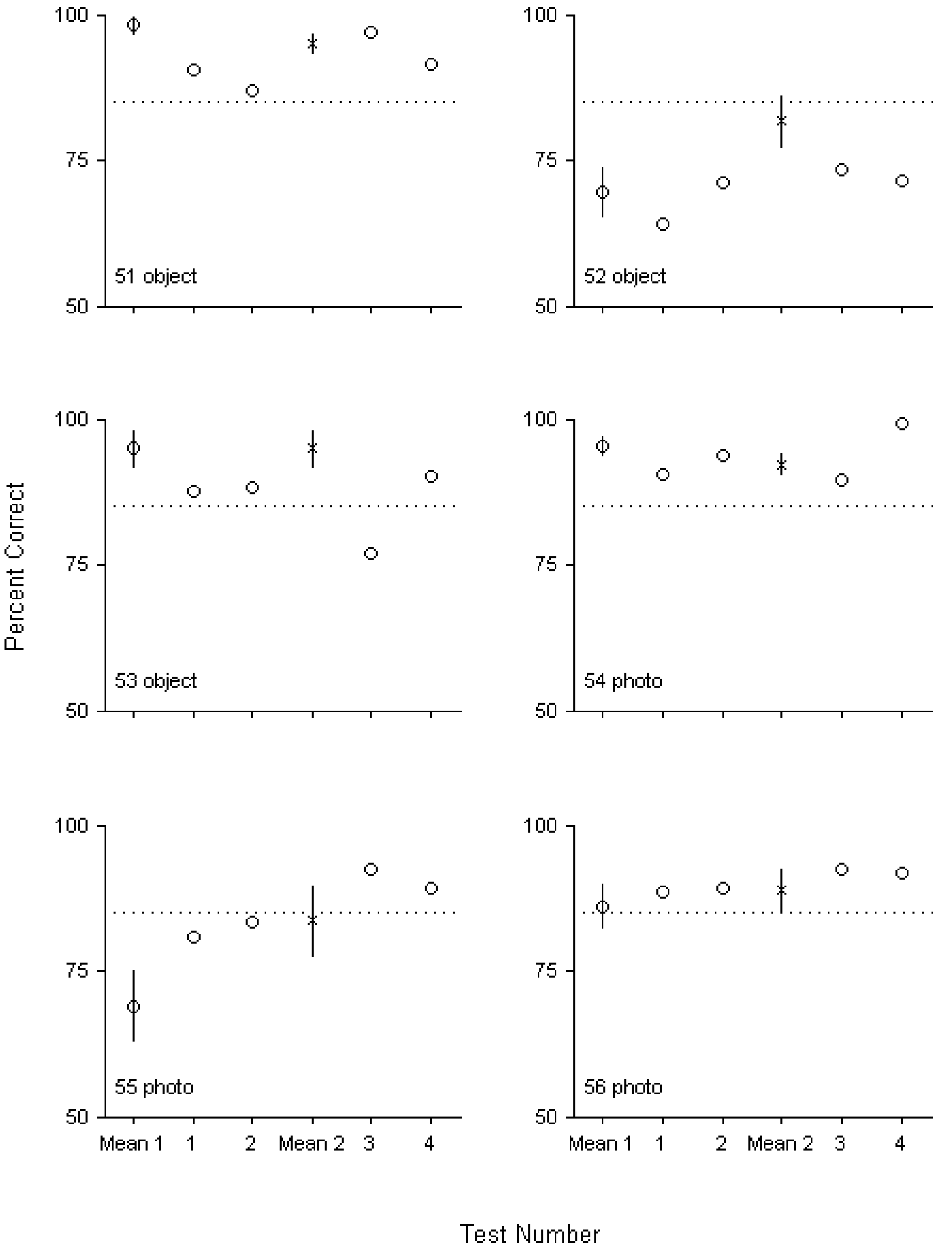


Figure 6.4. Mean percentage correct over the last five training sessions and the data from the test sessions (1-4) for all six hens (Condition 1). The first mean is the last five sessions before the first test session, and the second mean is the last five sessions before the final two test sessions. The vertical lines mark one standard deviation each side of the means and the horizontal dotted lines mark 85 % correct. Hens 51-53 were trained with the objects and tested with the photographs, and Hens 54-56 were trained with the photographs and tested with the objects.

($F(5,25) = 1.004$, $p>.05$, $\eta^2=.297$ (yet shows a moderate effect size)) across the data from all the hens from the test and training sessions.

Figures 6.5a and 6.5b presents the overall percentage correct, and the log *d* estimates for the last 40 training sessions of Condition 1, with the longer viewing tunnel, together with the data from the first 40 training sessions of Condition 2, when the viewing tunnel was shortened. It can be seen that accuracy for all hens was initially disrupted by shortening the viewing tunnel and bringing the stimuli closer to the hens. However, except for Hen 55, accuracy was regained to the same level shown with the long viewing tunnel within 18 training sessions. Hen 55's percentage correct never recovered to the same level as shown before the viewing tunnel was shortened, however her log *d* estimates were very similar across both conditions.

The left panel of Figures 6.6a and 6.6b shows the overall percentage correct over the training sessions, and the log *d* estimates, for all hens with the shorter viewing tunnel. Hen 52 was injured before test sessions began and as a result, missed five training sessions. There were 73 training sessions before the first test session. All hens (except Hen 55) were consistently responding above 85 % within 8 to 24 sessions and the log *d* plots showed discrimination remained high. The percent correct for these hens remained consistently above the 85 % over the remaining training sessions, except for Hen 53. For this hen accuracy was slightly more variable (range, 59 to 94 %) than that of the other hens and was frequently lower than 85 %. Hen 55's accuracy was never above 85 % over all 71 training sessions however, her accuracy was significantly above chance (that is, above 58 %). This is consistent with her results in Condition 1 where she required 85 training sessions before responding consistently above 85 %.

The right panel of Figures 6.6a and 6.6b presents the percentage correct for each stimulus, and the log *c* estimates during training in Condition 2. Hen 53 was slightly more accurate with the red object than the green object throughout training showing a bias to responding on the left key. Her accuracy with the green object was more variable (range, 35 to 96%) than with the red object. The rest of the hens showed no consistent biases of responding more to one key over the other.

Figure 6.7 shows the mean percentage correct over the last five training sessions and the data from the test sessions, for all six hens from Condition 2. Hen 55 completed test sessions even though her accuracy was not above 85 %, as she was responding above chance levels (58 %, binomial test). Hens 51, 52, 54 and 56's

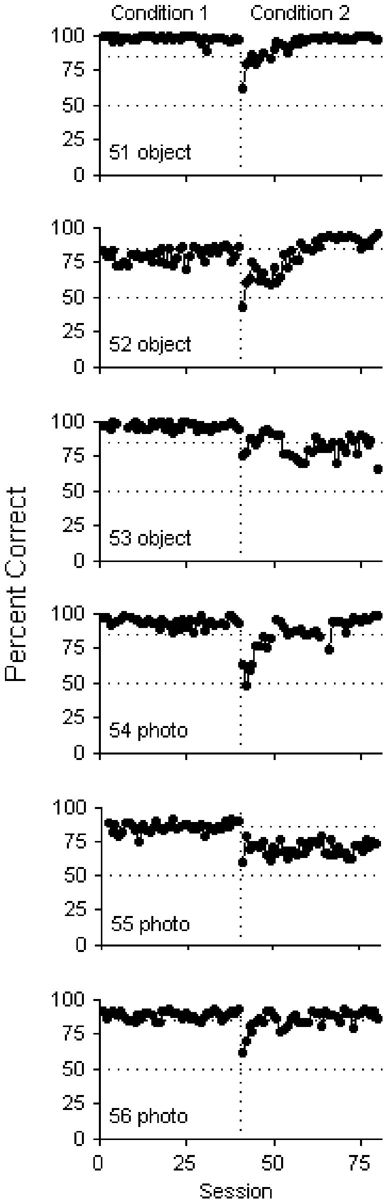


Figure 6.5a. Overall percentage correct over the last 40 training sessions in Condition 1 and the first 40 training sessions in Condition 2. The horizontal dotted lines mark 50 % and 85 % correct.

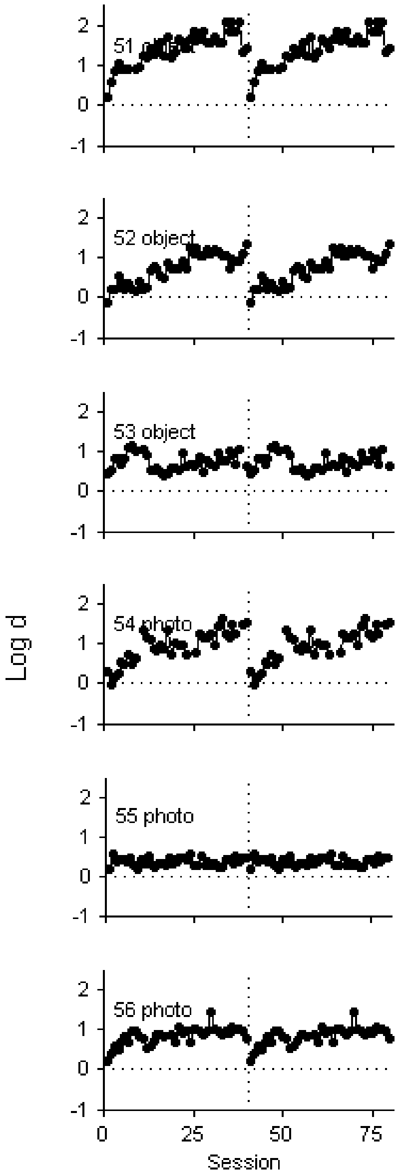


Figure 6.5b. Log *d* estimates over the last 40 training sessions in Condition 1 and the first 40 training sessions in Condition 2.

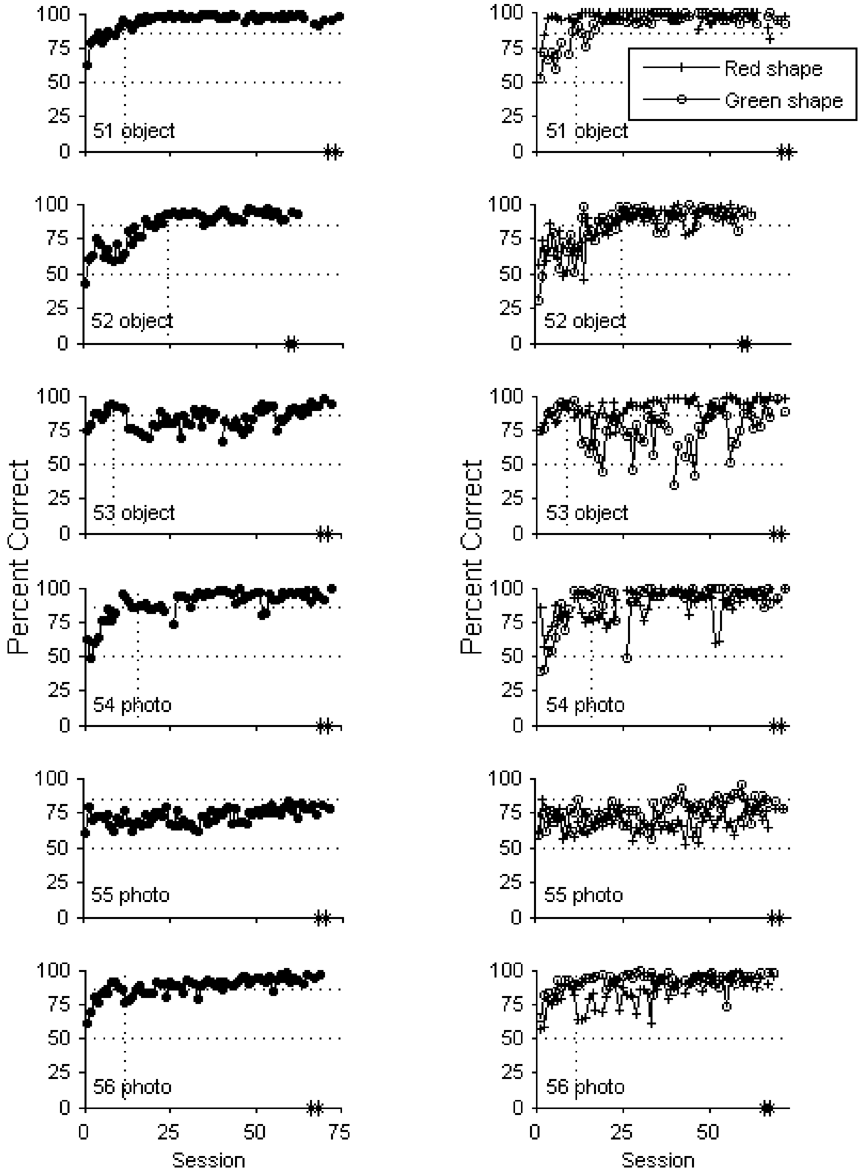


Figure 6.6a. The left panel shows the overall percentage correct over the training sessions with the red and green shapes across all sessions with the shortened viewing tunnel (Condition 2). The right panel shows the percentage correct for each of the stimuli. The vertical dotted line marks the point at which each hen had responded at or above 85 % for five sessions. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the locations of the test sessions.

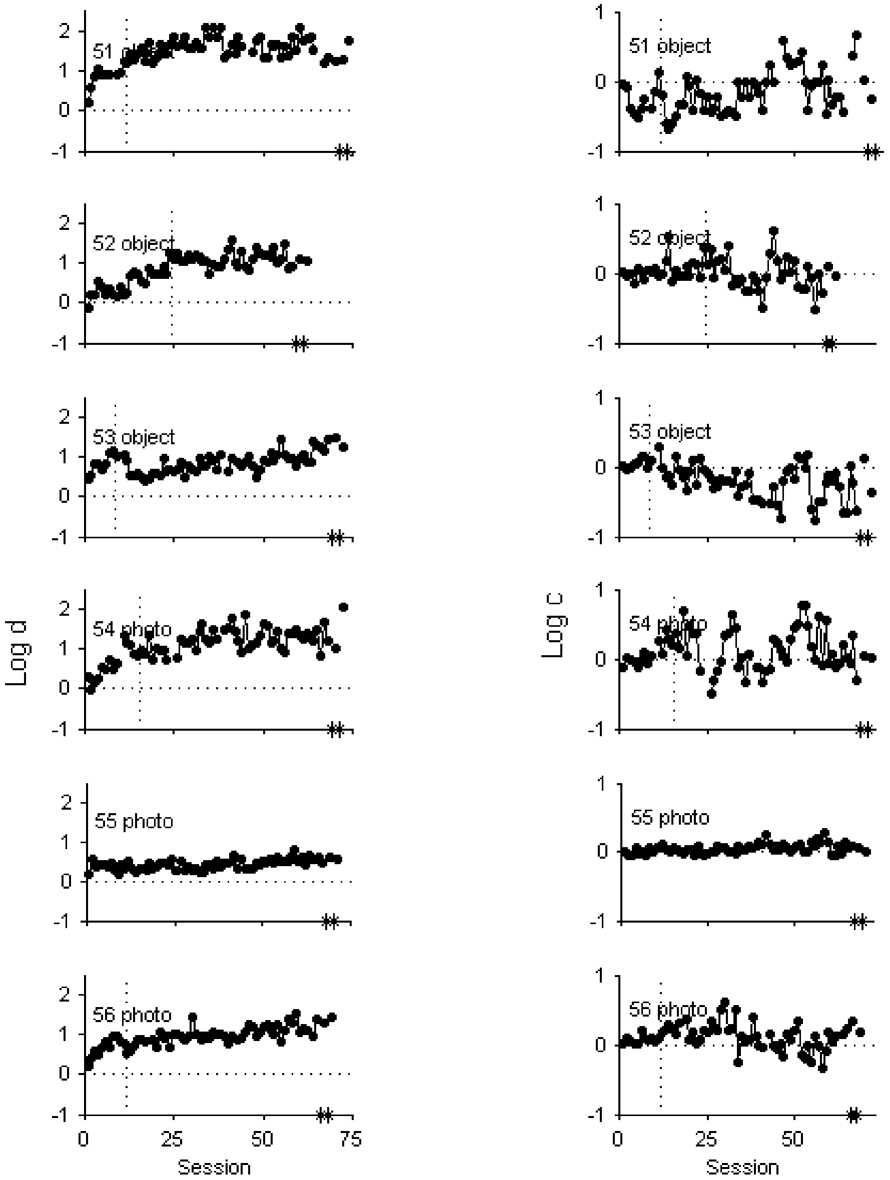


Figure 6.6b. The left panel shows the log *d* estimates over the training sessions with the red and green shapes across all sessions with the shortened viewing tunnel (Condition 2). The right panel shows the log *c* estimates across sessions. The vertical dotted line marks the point at which each hen had responded at or above 85 % for five sessions. The asterisks on the x axis represent the locations of the test sessions.

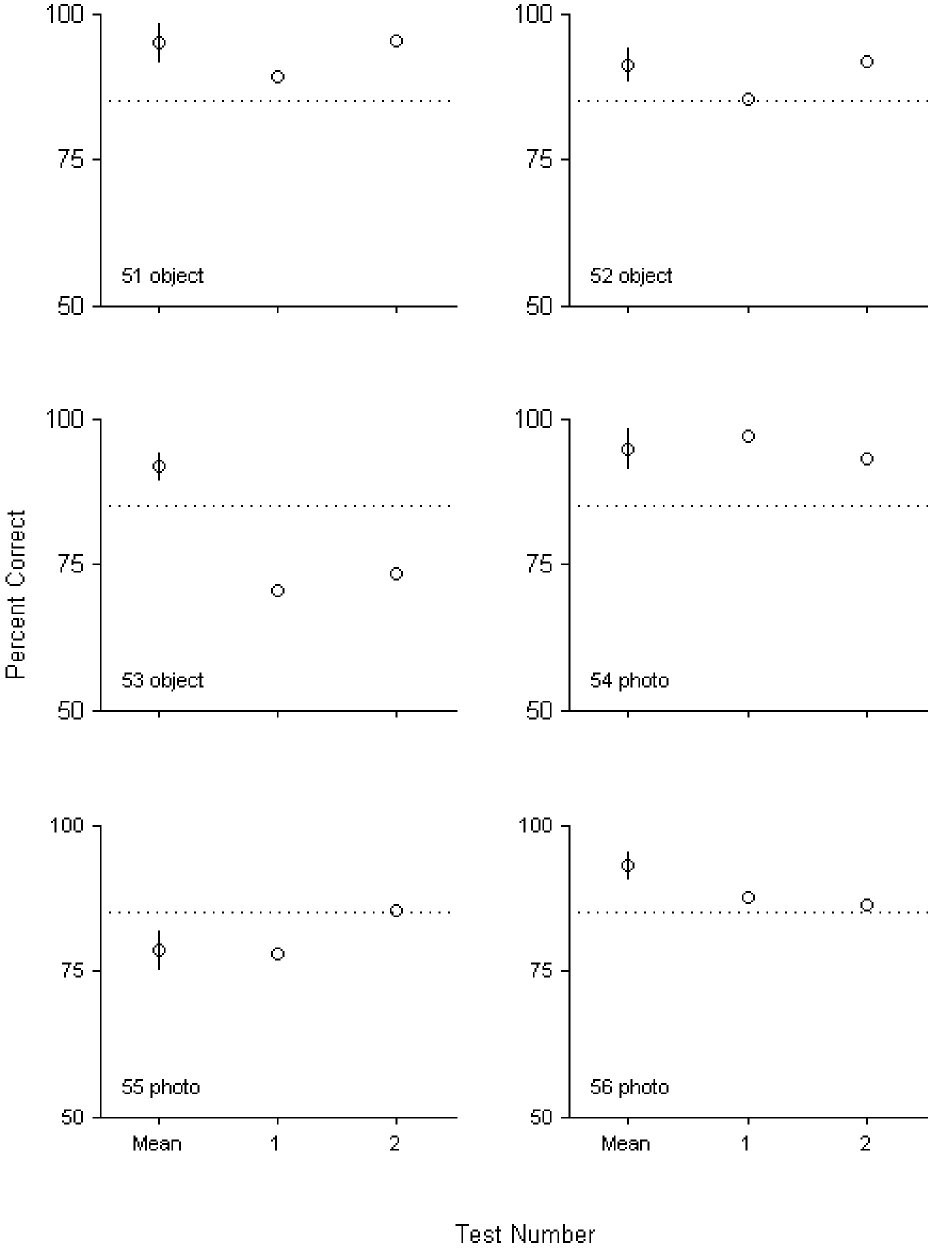


Figure 6.7. Mean percentage correct over the last five training sessions and the data from the test sessions for all six hens (Condition 2). The vertical lines mark one standard deviation each side of the mean and the horizontal dotted lines mark 85 % correct. Hens 51-53 were trained with the objects and tested with the photographs, and Hens 54-56 were trained with the photographs and tested with the objects.

accuracy during the test sessions remained high, and at or above 85 % (range, 85 to 97 %). Hen 53's accuracy dropped to 70 and 73 % respectively in each test session. Hen 55's accuracy was never above 85 % during training sessions, however, her accuracy during test sessions was relatively high at 78 and 85 % respectively.

Generally, for all hens, accuracy remained high in the test sessions when they were tested with the alternative stimuli from the training stimuli. A repeated-measures ANOVA for the data from the training and the two test sessions was not statistically significant ($F(2,10) = 2.528$, $p>.05$, $\eta^2=.392$). A Theil test could not be conducted with these data as *n* was less than 4.

Figures 6.8a and 6.9b show the overall percentage correct and log *d* estimates across the training sessions from Condition 3 (green shapes) for all hens. Log *d* and percent correct follow the same pattern. Percentage correct for all hens, except Hen 55, was at or above 85 % before test sessions began. Hens 53 and 54 learned the discrimination quickly, and were responding above 85 % correct within 15 training sessions. Hens 51, 52 and 56 took somewhat longer to learn but were eventually responding above 85 % correct within 90, 79 and 107 training sessions respectively. Percentage correct for Hen 55 was never above 85 %, over all the 220 training sessions. However, her accuracy was significantly above chance (that is, above 58 % using the binomial test). For all hens, except Hen 53, accuracy was stable before and after testing sessions, showing that the test sessions did not disrupt performance. Hen 53's accuracy was variable during the periods the test sessions occurred.

Figures 6.9a and 6.9b present the percentage correct for each stimulus, and the log *c* estimates over training sessions for Condition 3. Hens 51-53 were trained with the objects. Hen 51 showed higher accuracy with the triangular shape than with the rectangular shape up to session 100, but was slightly biased to responding on the key throughout all session. After session 100 she showed equally high responding to both stimuli. Hen 52 showed generally high accuracy with both stimuli, however, she showed a right key bias (associated with the rectangular shape) in the later training sessions. Hen 53 showed higher accuracy with the triangular shape than with the rectangular shape. Hens 54-56 were trained with the photographs of the objects and Hen 56's performance showed no consistent differences in accuracy and no consistent biases across the two stimuli or keys. Hens 54 and 55 were both more accurate with the rectangular shape than with the triangular shape, but showed no consistent biases.

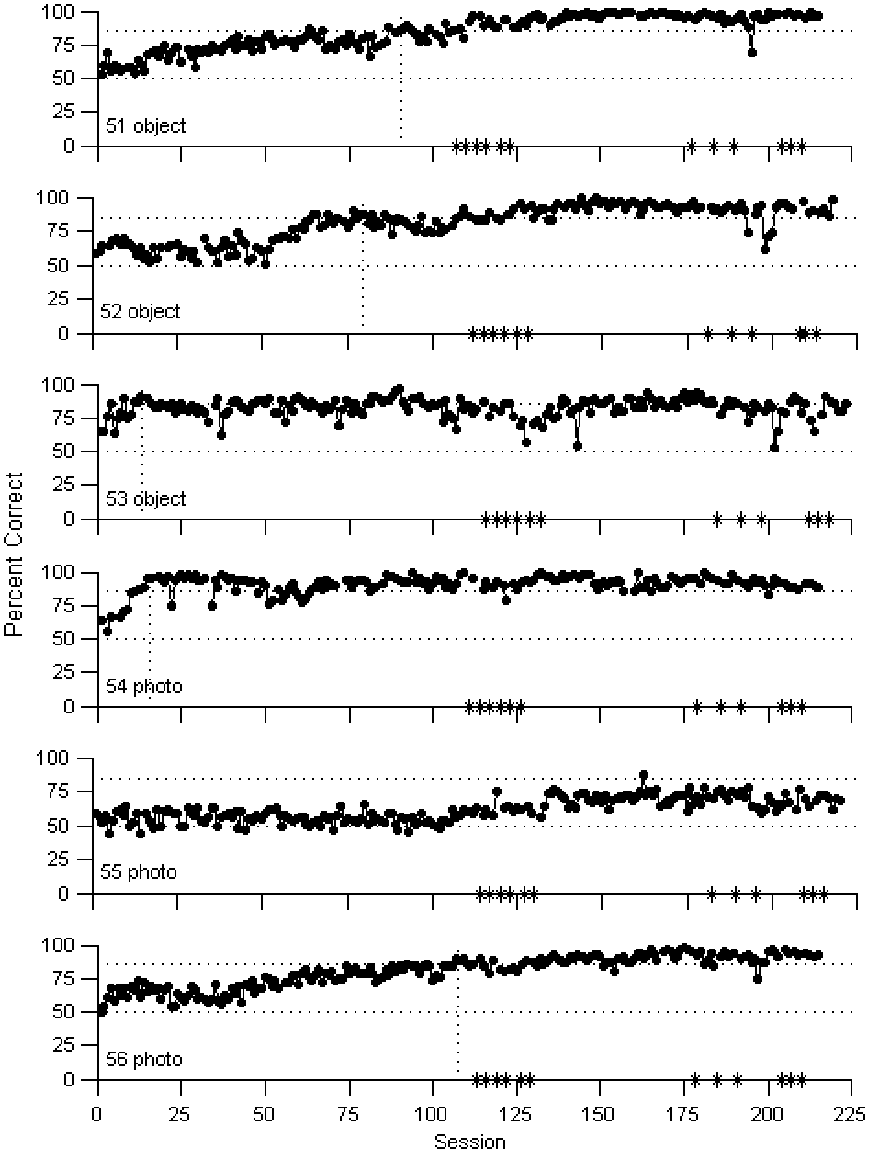


Figure 6.8a. Overall percentage correct with the green shapes (Condition 3) over the training sessions for all hens. The vertical dotted line marks the point at which each hen had responded at or above 85 % correct for five sessions. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the locations of the test sessions.

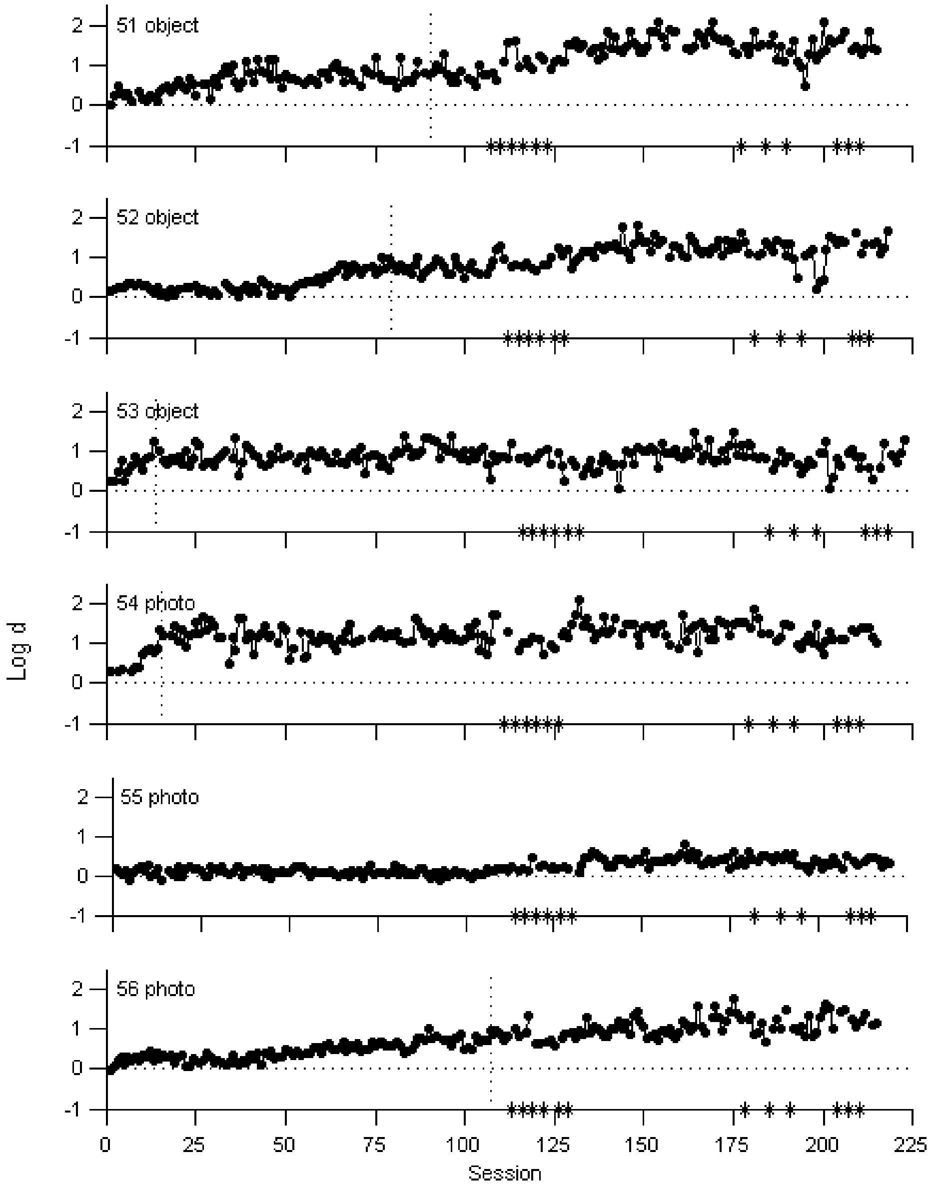


Figure 6.8b. Log *d* estimates over the training sessions for all hens. The vertical dotted line marks the point at which each hen had responded at or above 85 % correct for five sessions. The asterisks on the x axis represent the locations of the test sessions.

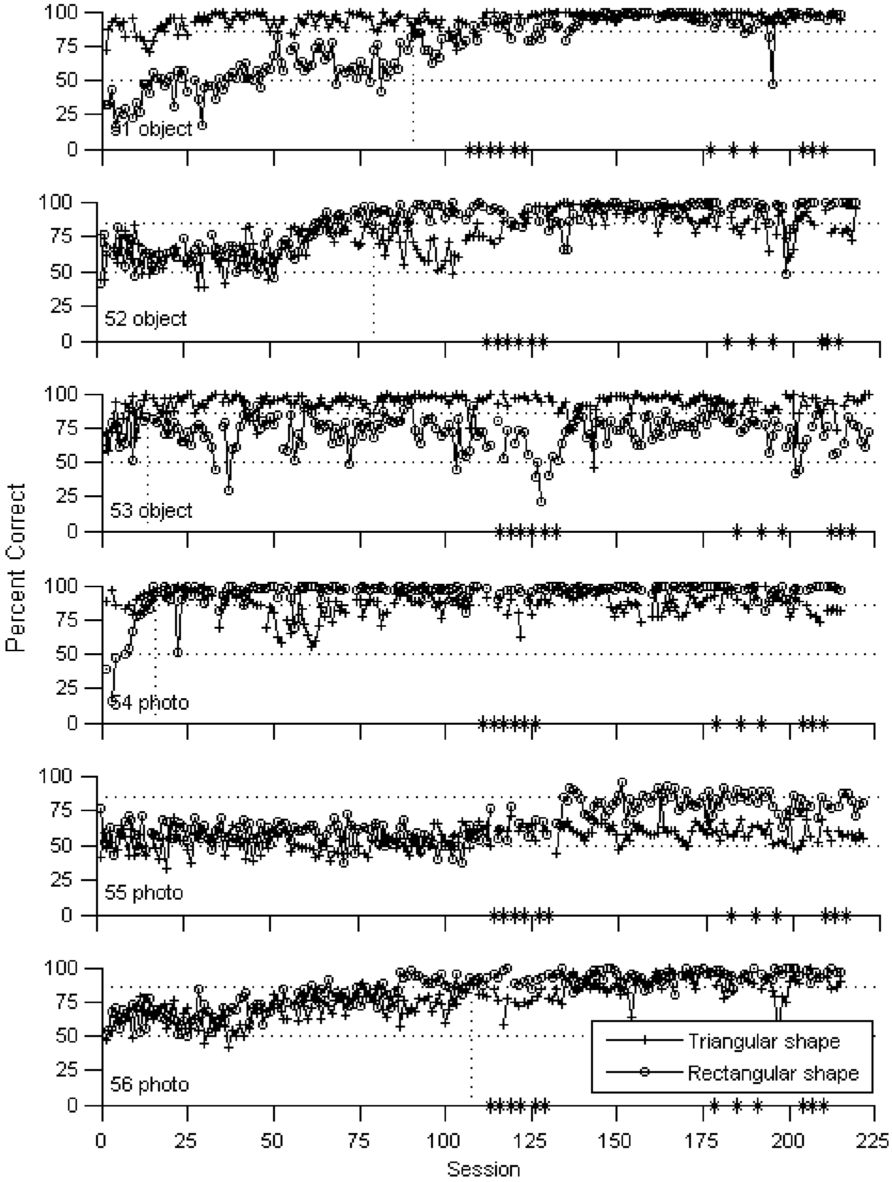


Figure 6.9a. Percentage correct for each stimulus over training sessions (Condition 3). The vertical dotted line marks the point at which each hen had responded at or above 85 % for five sessions. The horizontal dotted lines mark 50 and 85 %. The asterisks on the x axis represent the locations of the test sessions.

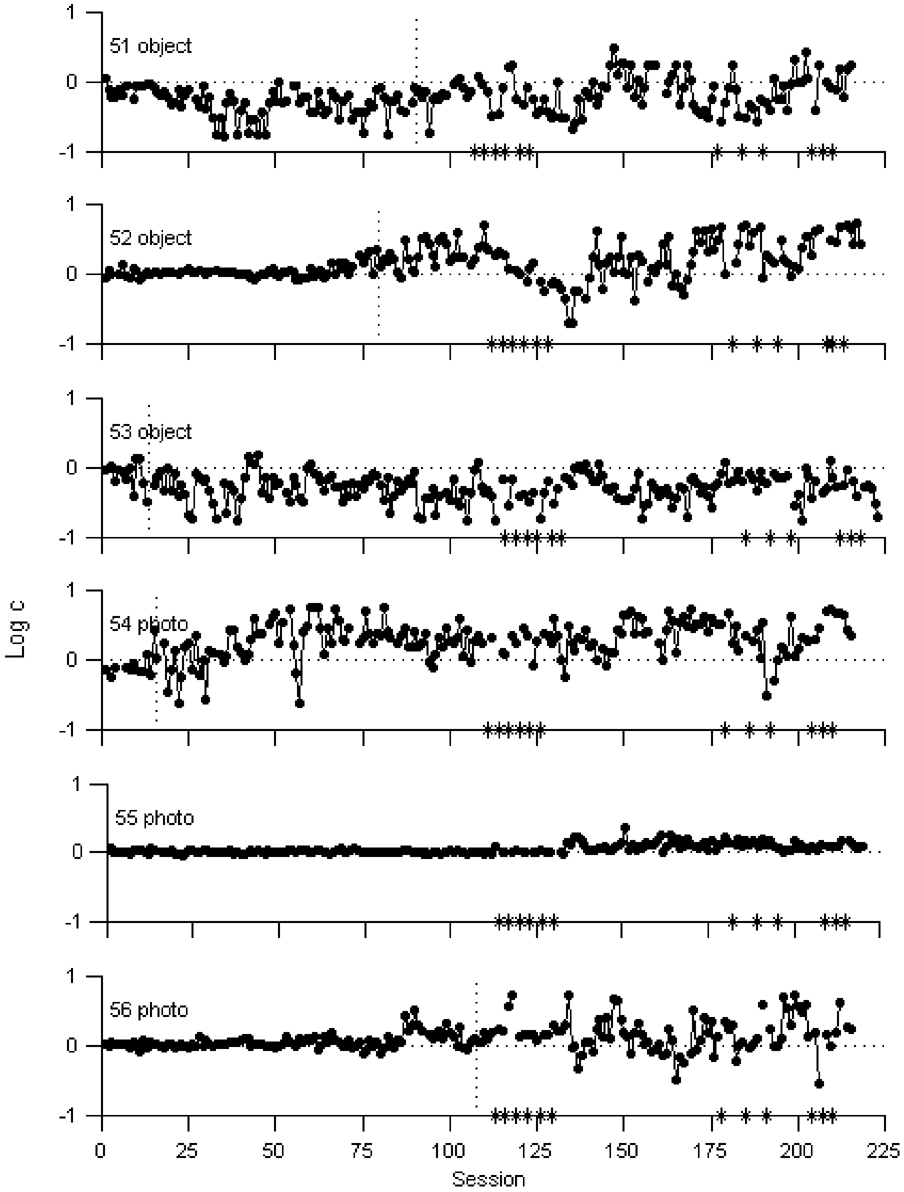


Figure 6.9b. Log *c* estimates over training sessions (Condition 3). The vertical dotted line marks the point at which each hen had responded at or above 85 % for five sessions. The asterisks on the x axis represent the locations of the test sessions.

Figure 6.10 presents the mean percentage correct and standard deviations over the last five training sessions, along with the mean percentage correct and standard deviations for each of the six types of test sessions. A repeated-measures ANOVA showed statistically significant differences over the data from the last five sessions of training and the six types of test sessions ($F(6,30) = 5.639$, $p<.05$, $\eta^2=.989$ (large effect size)).

It can be seen that during the transfer test sessions with the alternative stimuli (test session Types 1 and 2), all hens' accuracies, except for Hen 52, were higher (range, 54 to 94 %) when stimuli were presented in the training wedge (Type 1), than when the stimuli were presented in the non-training wedge (Type 2 - range, 40 to 76%). In fact, for all but one hen (Hen 52), all hens' accuracies decreased to around 50 % correct when the stimuli were presented in the non-training wedge, showing that transfer was based on cues available in the training wedge rather than the stimuli. In contrast to the other five hens, Hen 52 showed relatively high accuracy across both test session types (range, 62 to 76 %), indicating that she showed some degree of transfer to the stimulus, rather than control by some aspect of the training wedge used to present the stimuli.

The next two types of test sessions attempted to test for unintended sources of stimulus control by presenting the stimuli that the hens had been trained with in either the wedges that were not used during training (Type 3), or the wedges used during training, but with the edge of the wedge was masked (Type 4). When the stimuli that had been used in training were presented in the non-training wedges (Type 3), accuracy for Hens 51, 52 and 56 remained relatively high (range, 74 to 96%). However, accuracy for Hens 53, 54 and 55 was close to 50% (range, 46 to 60%), showing that these hens did not transfer their discrimination to the stimuli they had been trained to discriminate in a different wedge. When the edge of the wheel was masked, but all other factors remained the same as training (Type 4), three of the hens' accuracies (Hens 51, 53, and 55) were close to 50% (range, 48 to 60 %). The other three accuracies were relatively high. A paired-samples *t*-test showed that there was a significant difference in the data from these two types of test sessions ($t(5)=3.527$, $p < .05$, $d=1.446$ (moderate effect)). These findings indicate that these

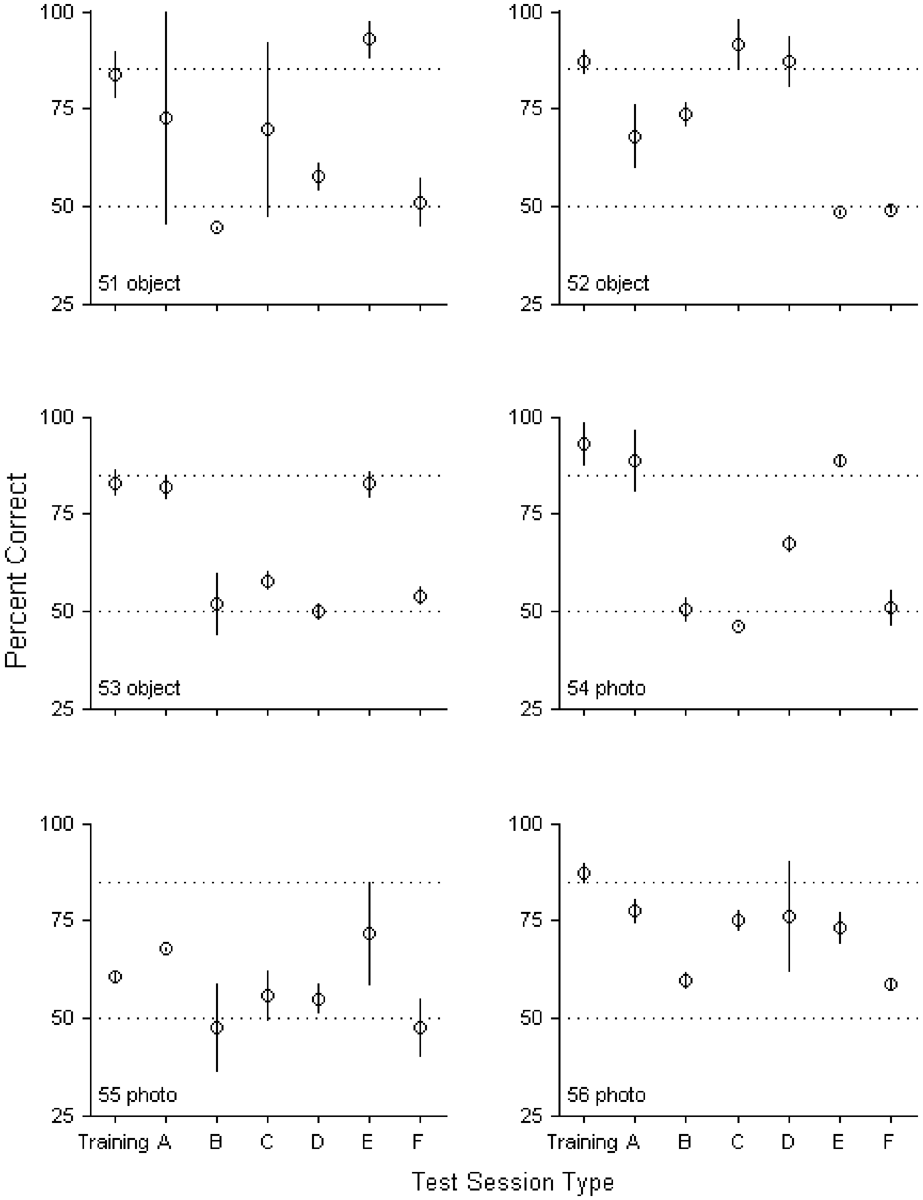


Figure 6.10. Mean percentage correct over the last five training sessions and the data from test session types 1-6 for all six hens (Condition 3). The vertical lines mark one standard deviation each side of the mean and the horizontal dotted lines mark 50 and 85 %. Hens 51-53 were trained with the objects and tested with the photographs, and Hens 54-56 were trained with the photographs and tested with the objects.

hens may have been using differences in the appearance of the particle board on the edges of the wheel in their discrimination, rather than the stimuli.

Stimulus control was assessed further in test session Types 5 and 6 by removing the stimuli. That is, the objects were removed or the photographs used were of the empty wedge. Five of the six hens (excluding Hen 52) showed relatively high accuracy (range, 63 to 96 %) when presented with the wedge used in training (Type 5 test), showing that, for these hens, discriminative performance was probably not based on the stimuli, but on the wedges that had been used during training. However, when the edge of the wheel was masked in test session Type 6, all hens accuracies were close to 50 % (range, 43 to 60 %). A paired-samples *t*-test showed there was a significant difference in the data from these two types of test sessions ($t(5) = -3.782$, $p < .05$; d=-1.897 (moderate effect)). In contrast to the other hens, Hen 52 showed low accuracy (range, 48 to 52 %) across both Type 5 and 6 tests. Thus, it appears that Hen 52 was the only hen that had learned to discriminate based on the stimuli, whereas the other five hens were discriminating based on features of the bottom edges of the wheel.

Discussion

One aim of this experiment was to assess if hens could learn to discriminate between two stimuli that differed in colour as well as shape, using the present procedure. The current data show that the hens were able to learn this discrimination. This finding suggests that the experimental equipment and the procedure could be used to assess discrimination. Patterson-Kane et al. (1997) similarly found that hens were able to discriminate between two real, or two videoed, stimuli when they were different colours. In the present experiment, the hens could have learned the discrimination based on colour, shape or a combination of the two. The simplest, and most likely, explanation is that the hens were using colour, however, these two factors cannot be separated here.

One factor the may have affected the hens' ability to discriminate is the length of the observing response. It was suggested previously, that increasing the time required for an observing response may have aided the acquisition of the discrimination task. The time that hens were in the presence of the stimulus could be taken as the time from when the beam was first broken to a key peck. It is possible that these times differed across hens and experiment and also that they may relate to whether or not the hens learned the discrimination. Thus, the median lengths of time from first beam break to key peck over the last five training sessions were calculated for Experiment 5 and the present experiment. It was found that median times ranged from 1.9 to 3.8 s (mean, 2.6 s) in Experiment 5, where all hens failed to learn the discrimination. In this experiment, where most hens showed high discrimination, the median times ranged from 1.9 to 4.8 s (mean, 2.6 s). Thus, median time spent in the presence of a stimulus prior to the keys being lit was similar across the two experiments. Therefore, the 0.5 s observing response used in these experiments was sufficient for the acquisition of this present discrimination. It must be noted that the hens also learned a colour discrimination in Experiment 2, but failed to learn a shape discrimination until an observing response was included in the procedure. In that experiment, the median length of time from first FR 5 key peck to key peck over the last five training session ranged from 3.7 to 6.9 s (mean, 4.6 s). As previously stated, DeMello, Foster and Temple (1993) found that hens' accuracy on a visual acuity task increased when the observing response requirements were increased. Similarly, White (1985) found that increasing an FR requirement increased pigeons' accuracy in a delayed-matching-to-sample (DMTS) task. Both studies show that increasing the response requirement can result in increased accuracy. It may be that increasing the observing response in this experiment (the minimum length of time that the beam needed to be broken) could increase hens' discriminative performance, particularly for more "difficult" discriminations, such as a discrimination based on shape rather than on colour. However, the effect that the length of the observing response has on discriminative performance was not examined in this thesis, but is an area that merits further study.

Once the hens had successfully learned to discriminate the differently coloured stimuli, the second aim was to examine whether hens transferred their discrimination of the 3D objects to their 2D photographs, and vice versa. The hens trained with the objects could discriminate between the photographs of the objects, and those hens trained with the photographs could discriminate between the objects themselves. As the hens' accuracy remained high during the test sessions, and the test stimuli were presented in the wedges that had not been used during training, this shows that the hens discriminative performance was indeed under control of the stimuli, rather than some other feature of the wedges.

Although it is possible that the hens may have learned to discriminate based on some unintended cue, it appears that this didn't occur with these stimuli. During the test sessions, the stimuli were presented in the wedges that were not used during training. If the hens had learned to discriminate based on some feature of the wedges used during training, accuracy would have decreased during the test sessions. Thus, it seems that the hens had learned to discriminate, and transferred that discrimination, based on the stimuli.

As previously mentioned, Watanabe (2000) states that when animals transfer their performance from object to pictures, this shows a degree of equivalence between these stimuli. This was shown in the present experiment where, at least with these differently coloured stimuli, hens responded to photographs in a similar way as to the objects. However, as stated earlier, it is likely that transfer with the coloured stimuli was under control of colour wavelength, rather than the shape of the object. In view of that, these findings do not provide conclusive evidence that the hens saw the images as a substitution for the object. This could be shown only when transfer occurred with stimuli that could be discriminated by shape alone.

The third aim of this experiment was to establish whether the discrimination was disrupted when the stimuli were moved closer to the hen. This change in viewing distance did disrupt performance for all hens. These findings are in line with those of Weaver (2000), where hens' percentage correct in a discrimination between two photographs of other hens' heads was suddenly reduced when the stimuli were moved closer to the hens. In that experiment, as in the present one, the discrimination was regained over several sessions. In the present experiment, all but one hen's accuracy quickly recovered. It is unclear why the hens' performances were disrupted by this change. It may have been a result of the way in which the hens viewed the stimuli, for example, they may have had to use more of their binocular field of vision with stimuli at closer distances. Alternatively, the stimuli may have appeared larger at the closer distances, changing the appearance of the stimuli. However, as viewing fixation points of the hens or apparent size of the stimuli were not controlled, no conclusions can be drawn about why accuracy was initially disrupted by this change.

The fourth aim was to establish whether the hens could learn to discriminate between two shapes when they were the same colour and presented at the shorter viewing distance. It does not appear that having previous experience with these shapes (albeit different colours) aided in the hens learning the discrimination when the same shapes were the same colour. That is, four of the six hens took longer to learn the discrimination when the shapes were the same colour, than when they were different colours. In Condition 3, one might expect better responding to the triangular shape than the rectangular shape, as this stimulus remained constant over all three condition, whereas the rectangular shape was changed from a red stimulus in Conditions 1 and 2 to a red stimulus in Condition 3. However, only one of the hens (53) showed higher accuracy with the triangular stimulus over the later training trials.

Testing in Condition 3 showed that, for most hens, accuracy remained high only when the alternative stimuli were presented in the wedges that were used during training. If the hens had been using the stimuli to transfer their discrimination, it would be expected that they should also show transfer to the stimuli when they were presented in the wedges that had not been used in training, as had been shown in the previous conditions; yet this did not occur in Condition 3. As a result, further tests were conducted with the stimuli the hens had originally trained to discriminate, to determine what factors were controlling discriminative behaviour. If the hens' discriminations during training were under control of the stimuli, it would be expected that accuracy would remain high regardless of what section of the wheel was used to present the stimuli, and regardless of whether the edge of the wheel was masked or not. However, it was found that three of the hens' accuracies were close to 50 % when the training stimuli were presented in non-training wedges, and three of the hens' (not necessarily the same hens) accuracies were close to 50 % when the edge of the wedge used in training was masked. These findings imply that these hens were not discriminating based on the stimuli, but rather some feature of the wedge sections of the wheel used to present the stimuli.

To further test this possibility, more test sessions were conducted in which the hens were presented with the wedges used in training, but no stimuli were presented. If accuracy decreased to chance, this would imply that the hens were discriminating based on the stimuli. If, however, there was still some degree of discrimination with no stimulus presentation, this would show that the hens had learned to use some unintended feature of the wheel in their discrimination. It was found that all, but one, of the hens still showed quite high accuracy when the stimuli were removed. Yet, when the edge of the wheel was masked, all hens were responding at chance. These findings show that five of the six hens were indeed discriminating based on other unintended and quite subtle features of the experimental equipment.

In light of these findings, the equipment was modified in the following experiment so that the edges of the wheel were obscured from the hens view, and all eight of the wedges were used in training and test sessions (rather than the four currently used) to reduce the probability of the hens' learning to discriminate on subtle cues that may be present in the wedges, rather than on the stimuli.

EXPERIMENT 7

The findings from the first two conditions of the previous experiment show that hens are able to distinguish between differently coloured stimuli, and that this discrimination will transfer to alternative stimuli (i.e., pictures, if trained with objects, or objects, if trained with pictures). The findings from Condition 3 of the previous experiment show that some hens' discrimination in this condition were under the control of small unintended features of the experimental equipment, such as wood grain patterns visible on the edge of the wheel, rather than the green shapes used as stimuli. As a result, no conclusions could be drawn about transfer from a 3D photograph to a 2D object with these hens. Thus, the present experiment aimed to remove any possibility of the hens learning to discriminate using cues other than the objects and so the experimental equipment was modified. The modifications included attaching a piece of Perspex to the bottom edge of the viewing tunnel during both training and test sessions, to prevent hens learning to discriminate based on wood grain patterns on the edge of the wheel. Also, in an attempt to reduce the probability of hens' learning to use other aspects of the training wedges as discriminative cues, all eight wedges were used in both training and test sessions. As a result it was decided to train all hens with the objects and use photographs in the transfer tests. As the hens used in the previous experiments had a lot of experience and had learned to discriminate on unintended features of the equipment, rather than the stimuli, a new group of hens, with no previous experience discriminating objects or photographs, were used in the following experiment.

The objects used in this experiment were the same Lego™ shapes that had been used to produce the 2D stimuli in Experiment 3. However, it must be noted that the first experiments reported in this thesis were conducted concurrently with the 2D/3D series and that the present experiment had been completed before Experiment 3 had begun. These objects were specifically constructed so that they differed only in shape. Thus, all pairs of stimuli were the same colours and textures, and equal numbers of blocks were used to construct the shapes so that their sizes were approximately equal. Different stimuli were used in two of the conditions reported here, as some hens learned to discriminate the shapes in the earlier condition and thus a new discrimination was needed for later conditions.

There were three conditions in the following experiment. In Condition 1, transfer from objects to photographs was examined, and, in Condition 2, transfer from photographs to objects was examined. There have been many studies that have looked for transfer by birds from 3D objects or conspecifics to 2D photographs or moving images (e.g., Cabe, 1976; Dittrich et al., 2010; Lumsden, 1977; Watanabe, Lea & Dittrich, 1993). However, relatively few studies have assessed whether transfer occurs in the opposite direction, that is, from photographs to objects. Of those studies that have (e.g., Spetch & Friedman, 2006; Trillmich, 1976; Watanabe, 1993), accuracy or performance was generally lower for those birds trained with photographs and tested with the real objects (or conspecifics).

Another way to examine transfer would be to determine if a discrimination would be learned more quickly if the animals had had previous experience with the stimuli presented in a different form. For example, if hens learned to discriminate objects more quickly (in fewer trials) when they had previously learned to discriminate photographs of the same objects – this would suggest a degree of transfer. Therefore, in Condition 3, the hens were trained with the objects that had been presented to them as photographs during training in Condition 2. If the hens did learn the discrimination in fewer sessions, this would imply that there was some transfer of the previously learned discrimination. On the other hand, if the hens did not learn the discrimination more quickly, this would imply that the previous experience with an alternative form of the stimuli did not aid in their learning, and that the object discrimination was essentially a new discrimination for the hens.

Method

Subjects

Six flock-reared Brown Shaver-Starcross hens (numbered 551-556) served as subjects. At the beginning of the experiment, the hens were two years old, and had had some experience on simple schedules of reinforcement, but no experience discriminating objects or photographs. They were individually housed in metal cages (500-mm long × 510-mm wide × 420-mm high) in a ventilated room that was lit on a 12:12-h light:dark cycle with two 20-W long life bulbs. Grit and vitamins were supplied weekly and water was available ad lib within their cages. Throughout the experiment all hens had red fleshy combs suggesting good health. Each hen was weighed before each experimental session (approximately six days per week) and they were provided with supplementary feed (commercial laying pellets) if required, to maintain them at approximately 80 % (+/–5 %) of their free-feeding body weights.

Apparatus

The apparatus was the same as that used in Condition 3 of Experiment 6, except that a piece of white Perspex (10-mm high × 115-mm wide) was screwed into place along the bottom edge of the viewing tunnel to prevent the hens from viewing the particle board texture on the edge of the wheel. In addition, only objects or photographs were now placed in eight of the nine wedges, therefore, all hens were either trained with the objects or photographs.

The stimuli used in Condition 1 were red and white Lego™ shapes (shown in Figure 7.1). Each had a white square base with a red square block clipped onto it. On top of this was either a red circular block (termed shape 1) or a square shaped block with one corner missing (termed shape 2). Four of each block type were used and all measured 16-mm long × 16-mm wide × 31-mm high. Photographs of the objects were adjusted (using Paintshop Pro, Version 7) to be the same size as the objects. A chroma meter (Minolta, CS-100) was used to match the colour of the photographs to be the same as the colour of the objects.

Figure 7.1. Image of the red and white Lego™ shapes used in Condition 1.

In Condition 2, the stimuli were blue and white Lego™ shapes (shown in Figure 7.2). Each stimulus consisted of a blue rectangular block clipped on top of a white rectangular block. Four of these had a square blue block clipped in the middle of the rectangular shapes (termed the inverted T-shaped block) and four had two small blue rectangular blocks clipped to each end of the rectangular blocks (termed the U-shaped block). All shapes measured 31-mm long × 16-mm wide × 31-mm high.

Figure 7.2. Image of the blue and white Lego™ shapes

Procedure

Condition 1. Sessions were similar to the discrimination training sessions reported in Experiment 6. However, all hens were now trained with the objects as the target stimuli. In addition, all eight wedges were used in the training sessions in an attempt to make it more difficult for the hens to learn to use unintended particular features of the wedges as cues for discrimination rather than the target stimuli. A conditional discrimination procedure was used; when shape 1 was the stimulus, a peck to the left response key was correct, and similarly when shape 2 was the stimulus, a peck to the right response key was correct.

At the beginning of discrimination training, there were four training sessions where only the correct key was lit, after which both keys lights were lit. The VR requirement was increased over a period of 80 days until all hens were responding on a VR3 schedule of reinforcement. Decisions to increase the VR were based on a combination of relatively stable performance and the subjects having received approximately 30 reinforcers in less than 40 minutes during the experimental session. The hens were considered to have learned the discrimination when they were responding at, or above, 85 % for five, not necessarily consecutive, sessions. Test sessions began for all hens after they had received at least 152 training sessions.

Test sessions, Condition 1. Test sessions followed the same procedure as used in discrimination training except the photographs of the objects were presented to the hens. All hens completed three test sessions with the stimuli presented in the same wedges as used in training. In addition, all hens had at least one training session between each test session.

Condition 2. The procedure was the same as that used in Condition 1 except that photographs of the target stimuli were used in training. A conditional discrimination was used; when the U-shaped block was the stimulus, a peck to the left response key was correct, and similarly when the inverted T-shaped block was the stimulus, a peck to the right response key was correct. The VR requirement was increased over a period of 33 days until all hens were responding on a VR3 schedule of reinforcement. Test sessions began after the hens had received at least 120 training sessions.

Test sessions, Condition 2. Test sessions followed the same procedure as used in Condition 2 training sessions except the objects were presented to the hens. Those hens that were responding above 85 % correct (553, 555 and 556) received a total of three test sessions. These hens received at least one training session between each test session.

Condition 3. The procedure and stimuli were the same as that used in Condition 2 except that the objects themselves were used in the training sessions, rather than the photographs. As the hens had had previous experience with this procedure and stimuli, the VR requirement began at VR3.

Results

All hens were trained with the red and white Lego^{TM} objects and tested with the photographs of the objects in Condition 1. Data were discarded if an egg was laid or if there were equipment problems during experimental sessions. Figures 7.3a and 7.3b shows the overall percentage correct and log *d* estimates for all training sessions for Condition 1, for all hens. Log *d* and percent correct follow the same pattern. Hens 553, 555 and 556's percentages correct increased to at or above 85 % relatively quickly (after 20-38 training sessions) and remained relatively stable for all remaining training sessions. Hen 554's percentage correct was at or above 85 % over five sessions after 95 training sessions, however, her performance was very variable and fell below 85 % in the later training sessions. Percentages correct for Hen 551 and 552 were never above 85 % even after 138 training sessions. Hen 551's performance was very variable (41 to 87 %), however, her responding was above chance levels before the test sessions began. In the five sessions before test sessions began there were a minimum number of 75 trials in a session, therefore, any percentage above 60 % was significantly different from chance at a .05 level of significant (binomial test).

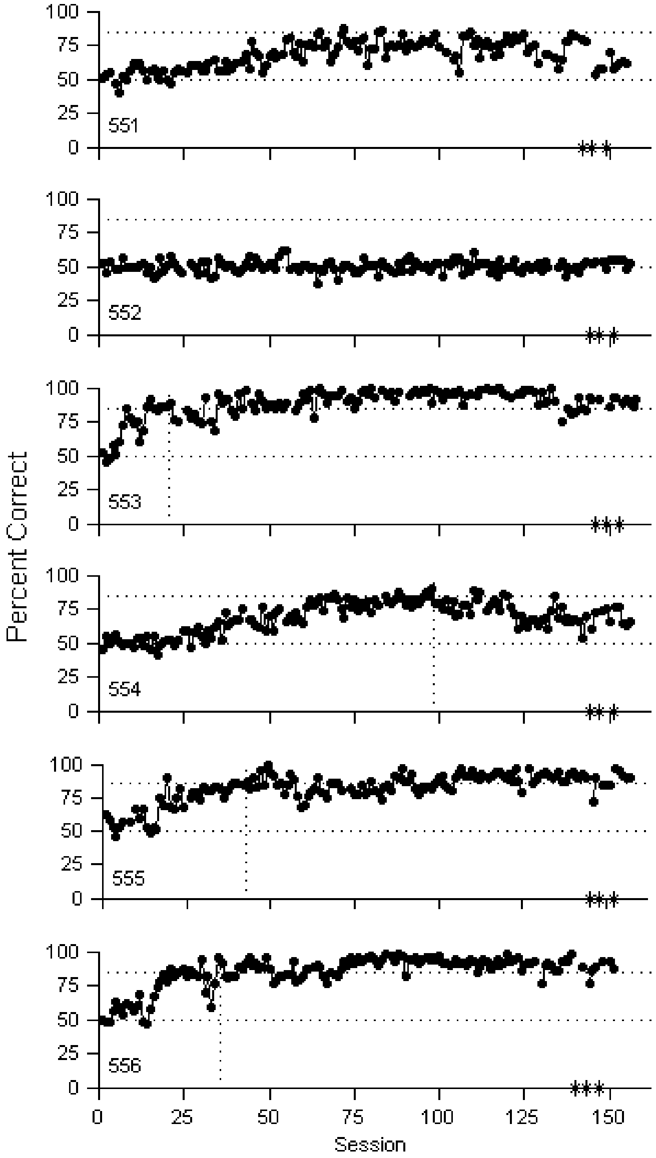


Figure 7.3a. Overall percentage correct over the training sessions from Condition 1 with the red and white Lego™ shapes for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the location of the test sessions.

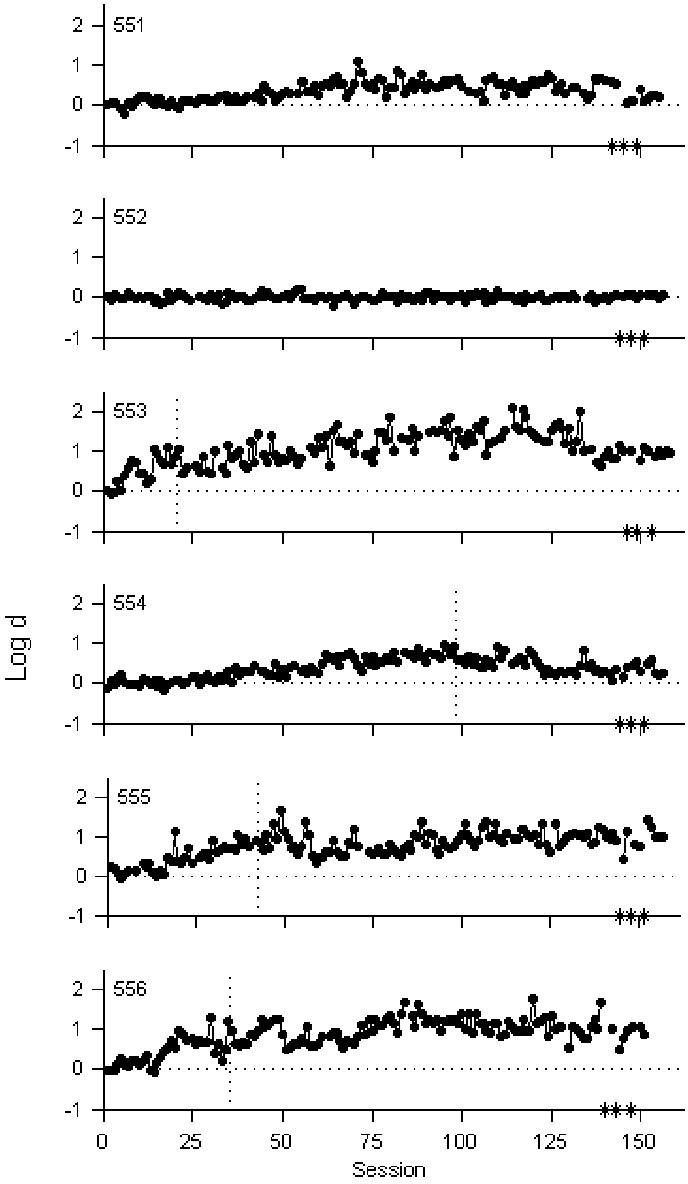


Figure 7.3b. Log *d* estimates over the training sessions from Condition 1 with the red and white Lego$^{\text{TM}}$ shapes for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the location of the test sessions.

Percentages correct for Hen 552 remained consistently at chance levels (50 %) over all the training sessions. Performance after test sessions remained consistent with performance before the test sessions for all hens, showing that the test sessions did not disrupt performance.

Figures 7.4a and 7.4b presents the percentages correct to each of the stimuli and log *c* estimates in Condition 1. There were no consistent differences in accuracy of responding across the two stimuli and no consistent biases for all hens.

Figure 7.5 shows the mean percentage correct over the last five training sessions and the data from the test sessions for all six hens for Condition 1. All hens completed test sessions, regardless of their performance during training. For all hens, except Hen 552, the first test session resulted in lower accuracy than their accuracy during the training sessions. For Hens 551-554, accuracy remained relatively low for the following two test sessions. Hens 555 and 556's percentages correct increased over the subsequent test sessions. A repeated-measures ANOVA showed there was a statistically significant difference across the data from the test and training sessions ($F(3,15) = 5.795$, $p<.05$, $\eta^2=.537$ (moderate effect size)). Paired sample *t*-tests show that there was a significant difference between the training data and data from the first and third test sessions ($t(5) = 2.737$, $p < .05$, $d=1.331$ (moderate effect); $t(5) = 4.572$, $p < .05$, $d=2.704$ (large effect) respectively). However, there was no significant difference between the data from the second test session and training data ($t(5) = 2.504$, $p > .05$, $d=1.044$). Paired-sample *t*-tests showed no significant differences in percentage correct between any pair of test sessions (test 1 and test 2, $t(5) = -1.036$, $p > .05$, $d=-.510$; test 1 and test 3, $t(5) = -.558$, $p > .05$ $d=-.245$; test 2 and test 3, $t(5) = 1.175$, $p < .05$ $d =.583$). Only Hen 551's data showed a significant trend (Theil test, C = 6, $n= 4$, $p<.05$). There were no significant trends in the other hens data (Theil test, C = 2, 0, 4, 2 and 0 for Hens 551 to 556 respectively, $n= 4$, $p>.05$

Figures 7.6a and 7.6b show the overall percentage correct across all training sessions and the log *d* estimates with the blue and white Lego^{TM} photographs in Condition 2. Percentage correct and Log *d* follow the same pattern. Hens 553, 555 and 556 required 37, 60 and 30 sessions respectively to reach at or above 85 % correct for at least five sessions. Performance for these hens remained high for the remaining training sessions. Hens 551, 552 and 554's percentages correct remained

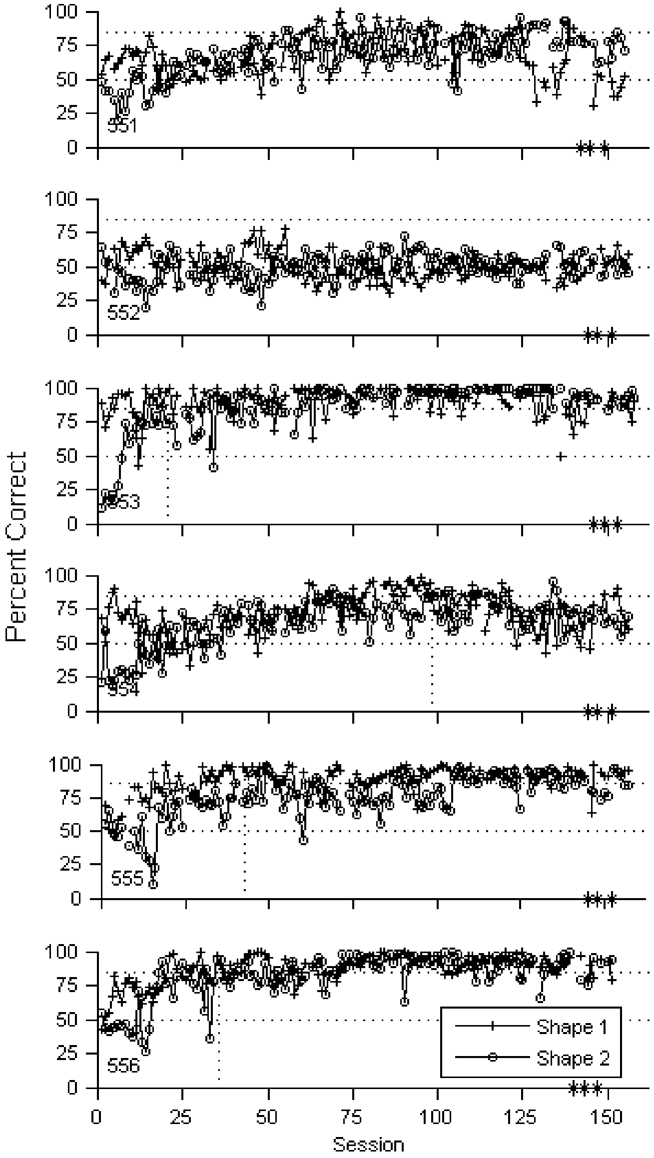


Figure 7.4a. Percentage correct towards each of the stimuli over training sessions with the red and white Lego^{TM} shapes (Condition 1). The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the location of the test sessions.

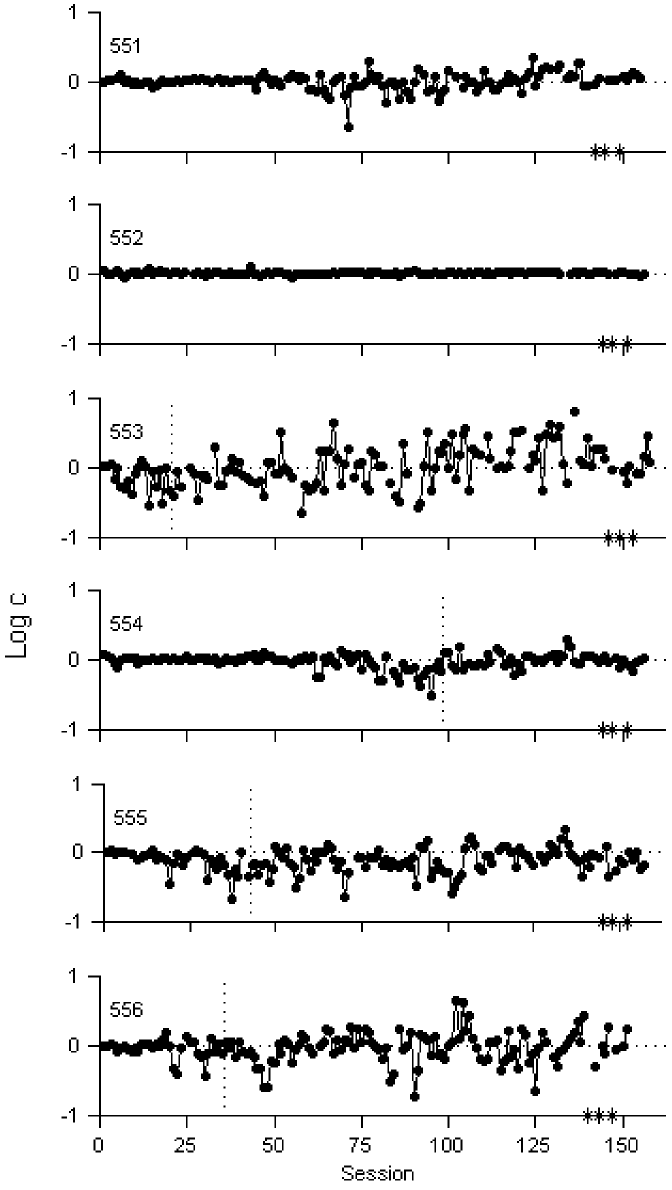


Figure 7.4b. Log *c* estimates over training sessions with the red and white Lego™ shapes (Condition 1). The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the location of the test sessions.

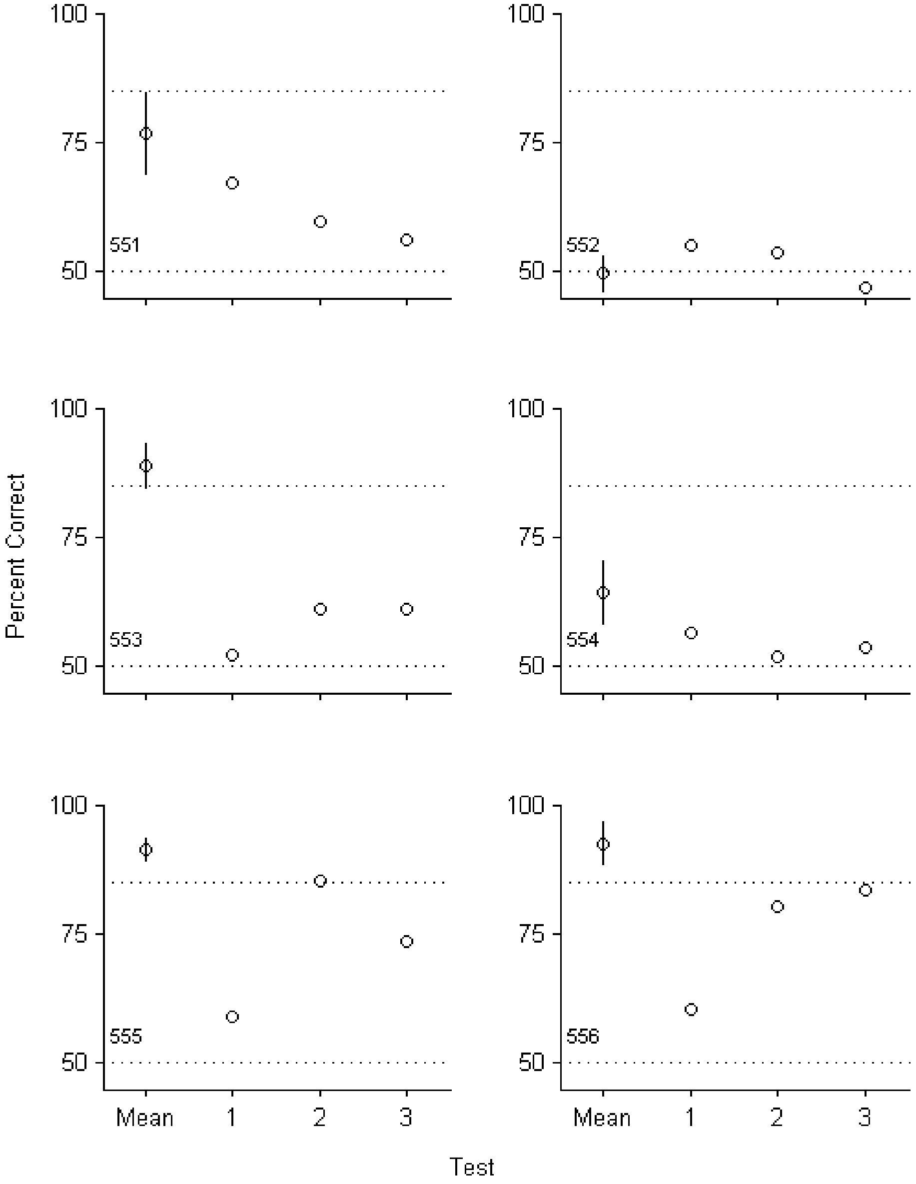


Figure 7.5. Mean percentage correct over the last five training sessions with the red and white Lego^{TM} objects and the data from test sessions 1-3 to the photographs (Condition 1). The training sessions were with the objects, and the three test sessions were with the photographs. The vertical lines mark one standard deviation each side of the mean and the horizontal dotted lines mark 50 % and 85 %.

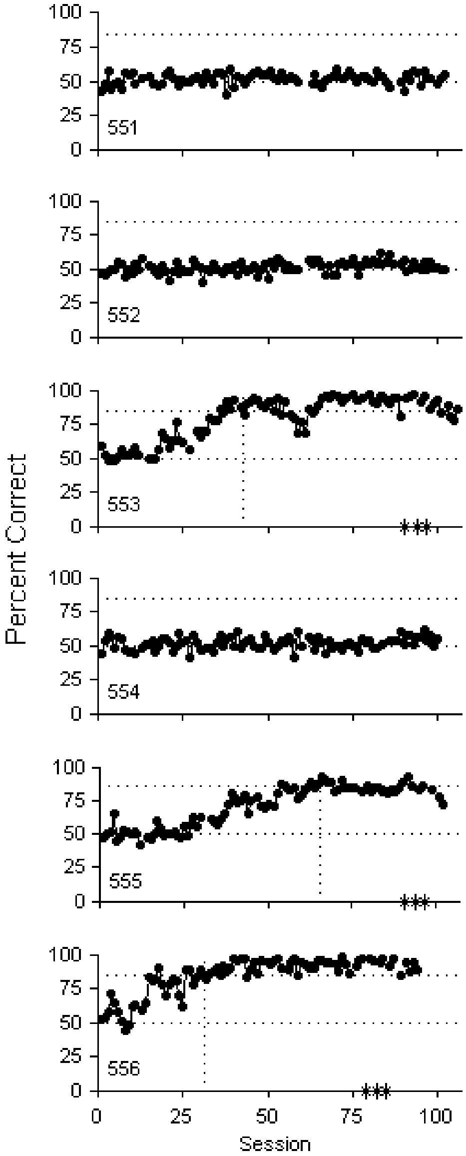


Figure 7.6a. Overall percentage correct over the training sessions with the blue and white Lego^{TM} photographs (Condition 2), for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the location of the test sessions.

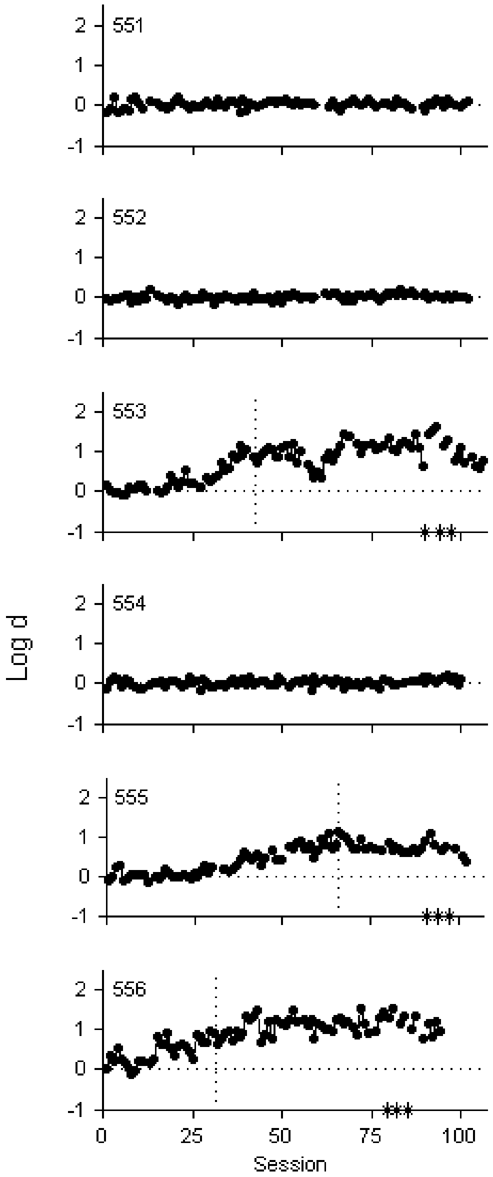


Figure 7.6b. Log *d* estimates plotted against training sessions with the blue and white Lego^{TM} photographs (Condition 2), for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the location of the test sessions.

at chance levels (50 %) after 115 training sessions. This is consistent with their performances in Condition 1, where Hen 551 and 552 never achieved 85 % correct and Hen 554 required more training than the other hens to reach 85 % correct.

Figures 7.7a and 7.7b presents the percentages correct to each of the stimuli, and log *c* estimates for all hens in Condition 2. No hen showed biases responding to one key over the other. Hen 551 showed higher accuracy with the photograph of the inverted T-shaped block (range, 66 to 83 % over the last five training sessions) than with the photograph of the U-shaped block (range, 21 to 39 %). Conversely, Hen 554 showed higher accuracy with the photograph of the U-shaped block (range, 66 to 81 % over the last five training sessions) than with the photograph of the inverted T-shaped block (range, 31 to 47 %). Hens 553, 555 and 556 show no consistent differences in responding across the two stimuli.

Figure 7.8 shows the mean percentage correct over the last five training sessions and the data from the test sessions for those hens whose performance was above 85 % correct during the training sessions (Hens 553, 555 and 556). The hens had been trained to discriminate between the photographs of the objects, and were tested for transfer to the objects during the test sessions. For all hens, accuracy over all three test sessions was lower (range, 47 to 65 %) than their performance during the training sessions. Small sample size numbers means that a larger effect is required to be statistically significant. In this condition, there was a sample size of 3. However, statistical tests were still conducted for comparison back to earlier data sets. A repeated-measures ANOVA showed there was a statistically significant difference across the data from the test and training sessions ($F(3,6) = 24.661, p<.05, \eta^2=.925$ (large effect size)). Paired-sample *t*-tests show that there was a significant difference between the training data and data from the second and third test sessions ($t(2) = 5.797, p < .05, d=3.379$ (large effect); $t(2) = 24.026, p < .05, d=19.436$ (large effect), respectively). However, there was no significant difference between the data from the first test session and training ($t(2) = 4.223, p > .05, d=2.452$ (moderate effect)). There were no significant differences in percentage correct between any of the test sessions (test 1 and test 2, $t(2) = -1.024, p > .05, d=0$; test 1 and test 3, $t(2) = 1.074, p > .05, d=-.645$; test 2 and test 3, $t(2) = 2.563, p < .05, d=1.485$). Hen 555's data showed a significant trend (Theil test, C = 6, , $n= 4$, $p<.05$). There were no significant trends in the data for the other two hens (Theil test, C = 2 and 4 for Hen 553 and 556 respectively, $n= 4$, $p>.05$).

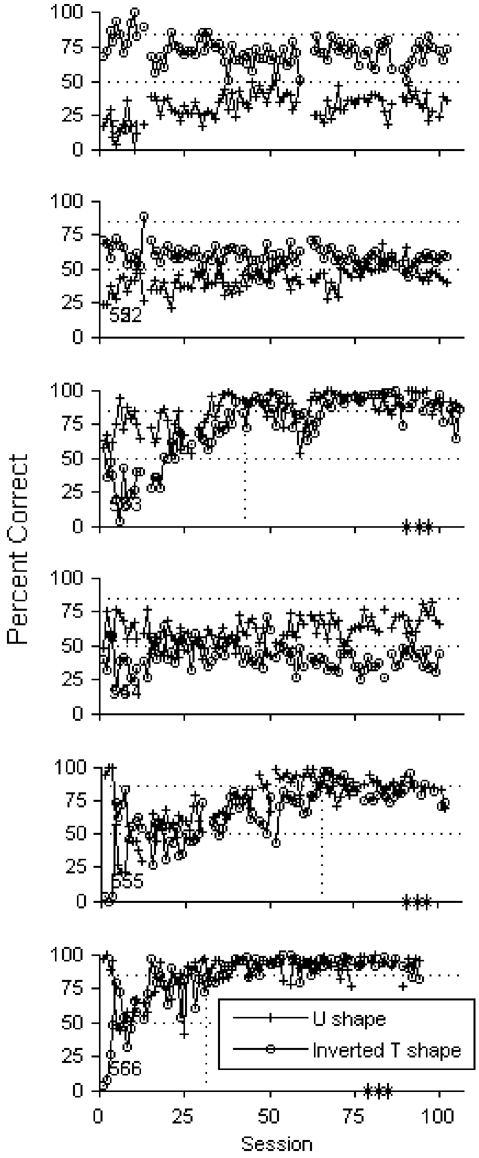


Figure 7.7a. Percentage correct towards each of the stimuli over training sessions with the blue and white Lego^{TM} photographs (Condition 2), for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the location of the test sessions.

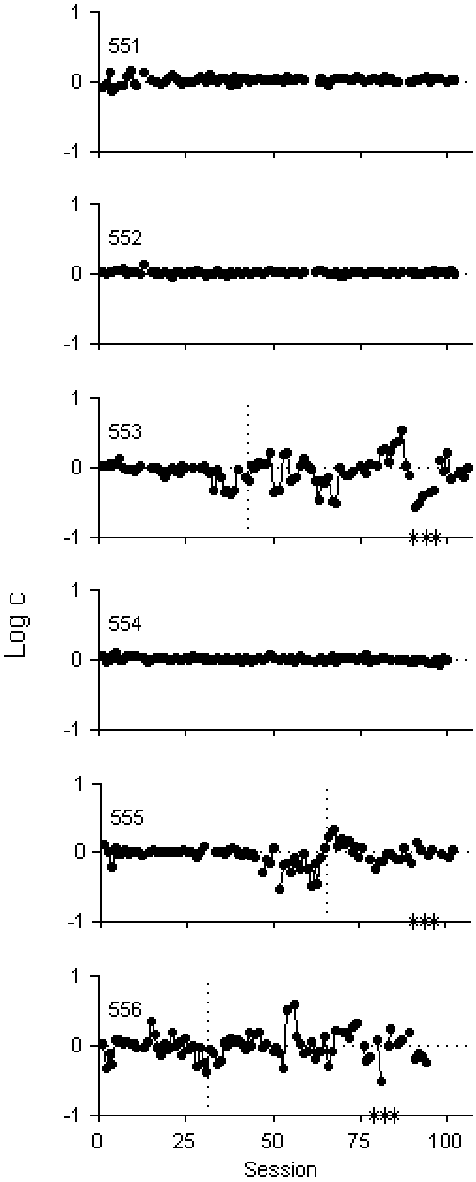


Figure 7.7b. Log *c* estimates over training sessions with the blue and white Lego™ photographs (Condition 2), for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the location of the test sessions.

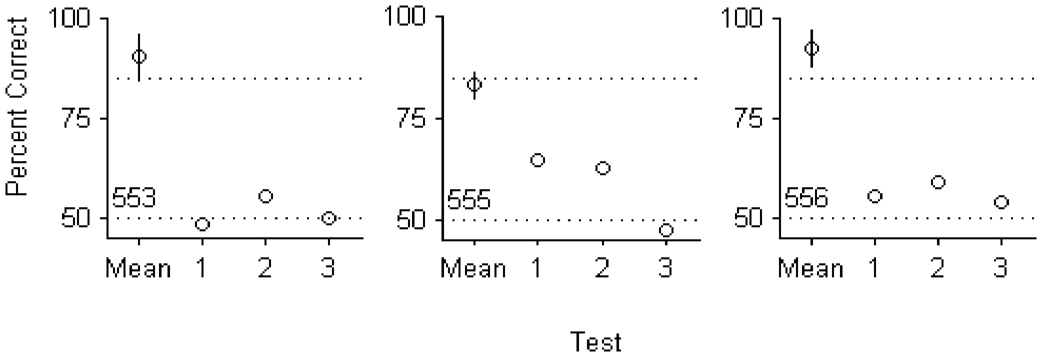


Figure 7.8. Mean percentage correct over the last five training sessions with the blue and white Lego™ photographs (Condition 2) and the data from the test sessions (1-3) with the objects for the three hens whose responding was above chance. The training sessions were with the photographs, and the three test sessions were with the objects. The vertical lines mark one standard deviation each side of the mean and the horizontal dotted lines mark 50 % and 85 %.

Figures 7.9a and 7.9b show the overall percentage correct and log *d* estimate for the last 20 training sessions with the photographs of the blue and white Lego ^{TM} blocks (Condition 2) and for the test sessions with the objects for the three hens that completed test sessions. Also shown are the percentages correct and log *d* estimates for all sessions with the blue and white Lego^{TM} objects (Condition 3). Log *d* and percentage correct follow the same pattern. For those hens that had learned to discriminate between the photographs in Condition 2 (Hens 553, 555, and 556), performance was initially disrupted when presented with the objects, showing no transfer of accuracy from the photographs to the objects. As the same wedges were used to present the stimuli, this suggests that the hens were using the stimuli to discriminate in Condition 2. In addition, the three test sessions with the photographs showed similar levels of accuracy to the initial training sessions with the objects. During Condition 3, Hens 553, 555 and 556 required 36, 62 and 43 sessions, respectively, to achieve 85 % correct or above over five sessions. Accuracy for these hens remained high for the remaining training sessions. Hens 551, 552 and 554's percentages correct all remained at chance levels (50 %) after 100 training sessions. These results are consistent with their performances in Condition 2.

Figures 7.10a and 7.10b present the percentages correct to each of the stimuli and the log *c* estimates in Condition 3. Only Hen 553 showed biased responding to one key over the other, that is, she responded more to the right key (associated with the inverted T-shape) than the left key. None of the other hens showed consisted biases in responding to a key. Hens 551-553 all showed slightly higher accuracy with the inverted T-shaped block than with the U-shaped block.

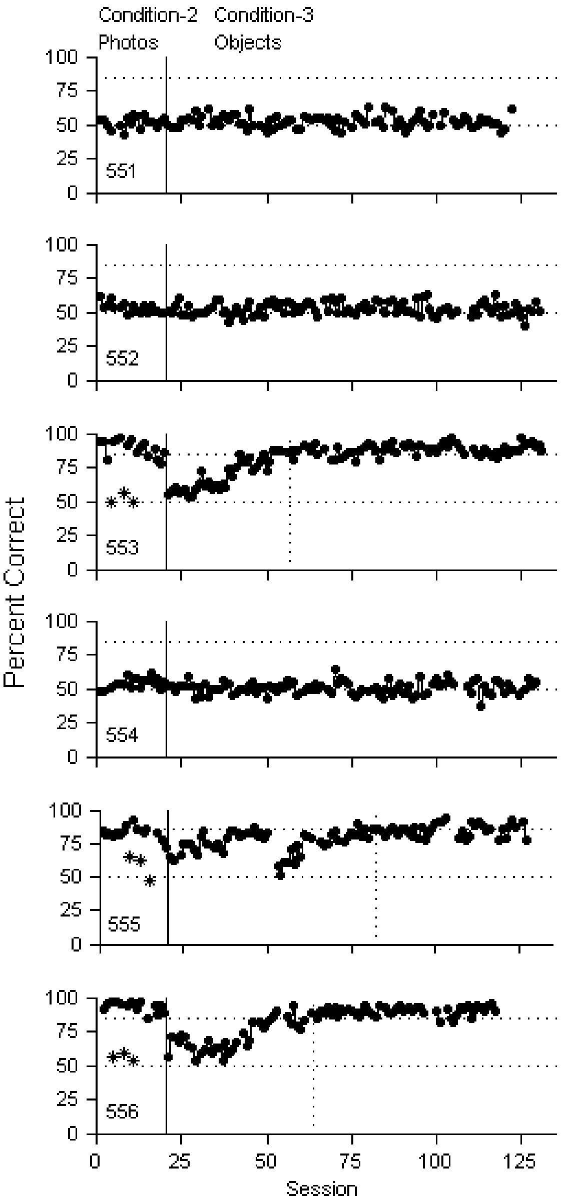


Figure 7.9a. Overall percentage correct for the last 20 sessions with the blue and white Lego™ photographs (Condition 2), and percentage correct with the blue and white Lego™ shapes (Condition 3) over all sessions. The vertical solid line marks the condition change to the objects. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks represent the test sessions with the objects.

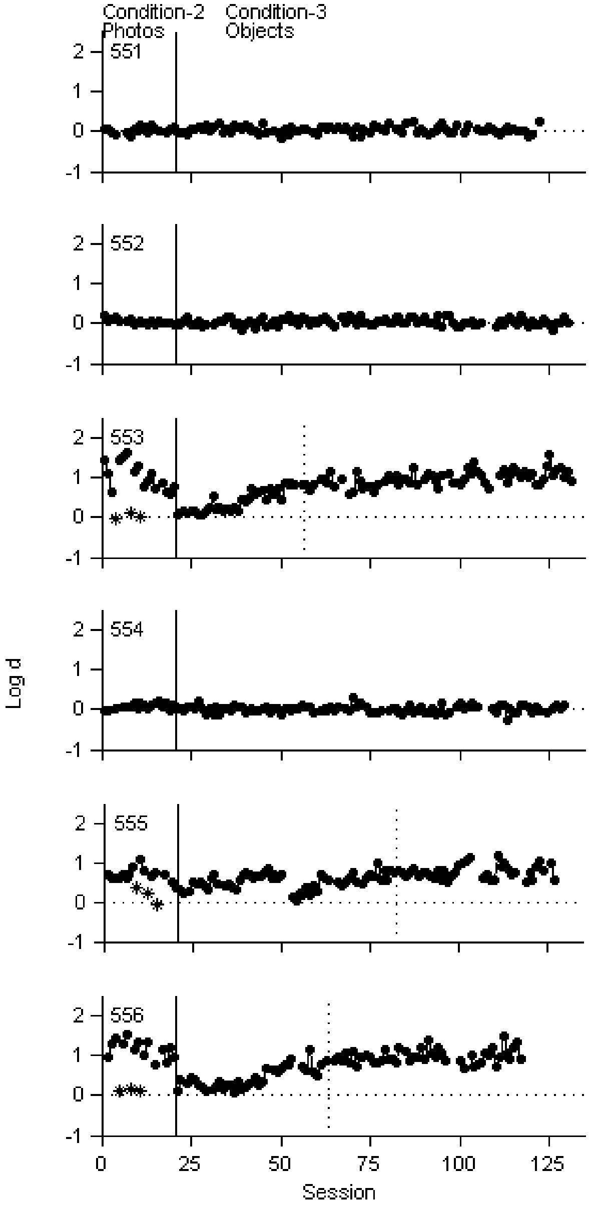


Figure 7.9b. Log *d* estimates for the last 20 sessions with the blue and white Lego™ photographs (Condition 2), and log *d* estimates with the blue and white Lego™ shapes (Condition 3) over all sessions. The vertical solid line marks the condition change to the objects. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks represent the test sessions with the objects.

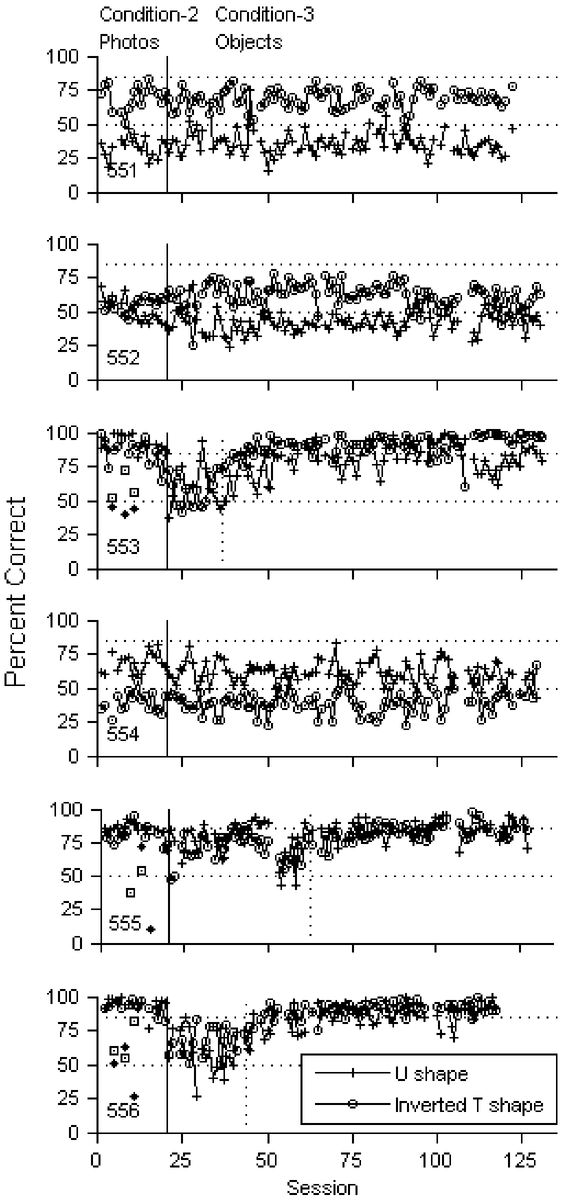


Figure 7.10a. Percentage correct towards each of the stimuli over training sessions for the last 20 sessions with the blue and white Lego^{TM} photographs, and percentage correct with the blue and white Lego^{TM} objects. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 % and 85 %.

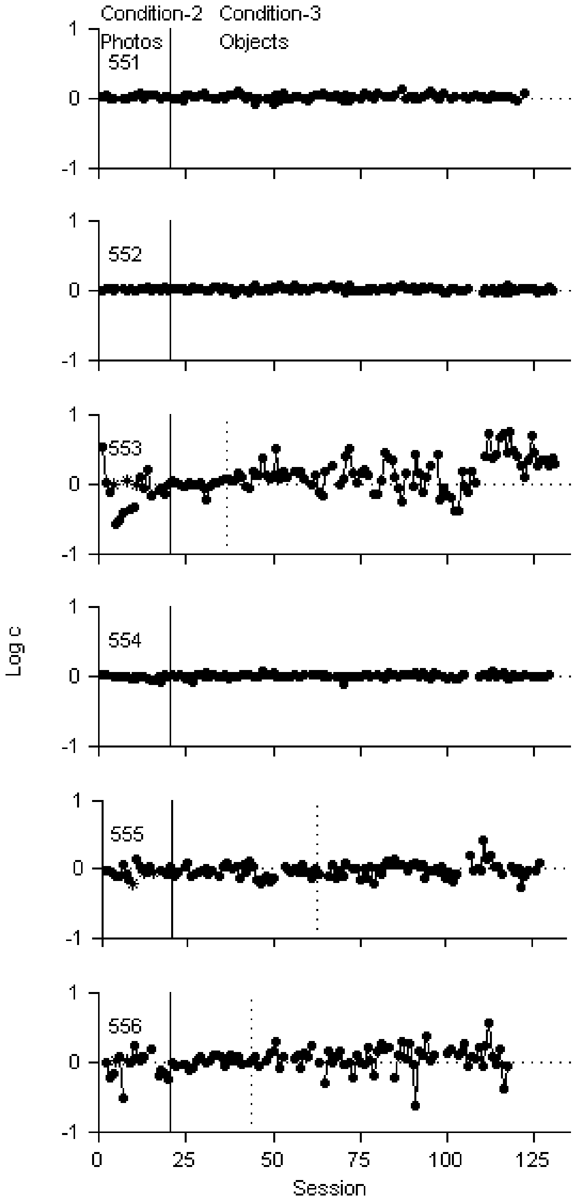


Figure 7.10b. Log *c* estimates over training sessions for the last 20 sessions with the blue and white Lego^{TM} photographs, and percentage correct with the blue and white Lego^{TM} objects. The vertical dotted line marks the fifth session that each hen responded at or above 85 %.

Discussion

The current data show that, for all conditions, some of the hens were able to learn the discrimination when the stimuli differed in shape only. However, some hens did not do so even after extensive training. In addition, none of the hens showed transfer to the alternative stimuli during the transfer tests. In Condition 1, although one hen showed relatively high accuracy (67 % correct) in the first test session, accuracy decreased in the next two sessions. If she had been using the stimuli during the transfer tests, accuracy should have remained high across all test sessions. In addition, while two hens had high accuracy in the last two test sessions in Condition 1, this could probably be attributed more to the hens learning a new discrimination, (as reinforcement was still available during test sessions) rather than evidence of transfer. However, this increase in accuracy over further test sessions was not shown in Condition 2, showing learning did not occur in that condition.

A similar increase in accuracy across test trials was shown in Spetch and Friedman (2006), who examined transfer of accuracy between objects and pictures. They trained pigeons to discriminate between either two objects or between two images of the objects and tested for transfer and reestablishment of discriminative performance by replacing pictures with objects and objects with pictures for each group of pigeons. They also reversed the discrimination for half of the pigeons so the stimulus associated with reinforcement was now the non-reinforced stimulus, and vice versa. The authors argue that the degree of difference between the two contingency groups during test trials established the pigeons' degree of transfer. Pigeons in the same-contingencies group showed higher overall transfer in both stimulus groups (picture to object and object to picture) than those in the reversed-contingencies group. However, this transfer was low in initial transfer trials. That is, while accuracy was high during the training trials (above 90 %), accuracy dropped for both groups in the initial (first 50 trials) transfer trials. For those pigeons in the reversed-contingencies group, accuracy dropped below 50 % for transfer to both real objects and pictures. For the same-contingencies group, accuracy dropped to about 55 % for those birds transferring to objects, and to about 70 % for those birds transferring to pictures. This decrement in performance in the initial transfer trials is similar to that shown in this experiment, where all hens had lower accuracy in the first test session than in the previous training sessions. Spetch and Friedman (2006) found that their pigeons' accuracy increased over further transfer trials. A similar increase in accuracy over repeated test sessions was also shown in Condition 1 of this experiment (object to photograph transfer) by Hens 555 and 556. As reinforcement was still available during test sessions in this study and in Spetch and Friedman's study, learning is a possible confound with both procedures. The increase in accuracy over further transfer trials could be the result of the birds learning a new discrimination, rather than transferring the previously learned discrimination. Consequently, it could be argued that the initial test session is the only one to truly indicate if transfer has occurred. Conditions 1 and 2 here show that, for those hens that learned to discriminate the stimuli, none transferred their performance to the real objects.

It appears as though these results show that the hens in this experiment were indeed using the stimuli in their original discrimination. That is, if the hens had learned to discriminate using extraneous cues, as they did in Condition 3 of Experiment 6, accuracy would have remained high during the test sessions. However, accuracy decreased for all hens, showing that the modifications made to the equipment and procedure successfully removed cues that the hens were using in Experiment 6 to learn a different, unintended, discrimination.

The results of Condition 1 indicate that testing hens who were not responding above 85 % is not necessary. In Condition 1, all hens received test sessions, regardless of whether their accuracy was at or above 85 %. However, analysis of Hen 552's results show that there was really no new information to be gained by testing a hen whose responding was not above chance levels. This hen was not doing the original discrimination and so there was no discrimination to transfer. Therefore, only those hens whose performances were above chance received test sessions in Condition 2.

It is possible that transfer from photographs to objects may be more difficult than transfer from objects to photographs. In Condition 2, accuracy in the initial test sessions was low for all three hens, showing that no transfer occurred. That is, these hens did not transfer their learned discrimination of the photographs to the objects themselves. In addition, their accuracy for the following two test sessions remained low, indicating that unlike Condition 1, no learning occurred. Spetch and Friedman's (2006) results also show lower overall accuracy for the group that were trained with the photographs and tested with the objects. The current findings are consistent with those of Spetch and Friedman, and support the idea that birds find it more difficult to transfer discrimination from photographs to objects compared to transfer from objects to photographs.

It was considered that previous experience with photographs may have an effect on how quickly the hens learned to discriminate the objects. That is, hens that had previous training with the photographs of the objects may learn to discriminate between the objects more quickly than they had learned with the photographs. However, the hens did not learn to discriminate the objects more quickly after having learned to discriminate the photographs, showing that these hens did not transfer their performances to the objects, and this was a new discrimination. These findings further suggest that photographs were not equivalent to object for these hens.

These same images of both sets of stimuli had also been used in Experiment 3. However, those hens never learned to discriminate between the images of the objects, and in the present experiment, only three of the six hens learned to discriminate between either the photographs or the objects. Although the stimuli appear to differ to the human eye, the hens had difficultly discriminating the stimuli. As previously mentioned, hens can do conditional discriminations (e.g., DeMello, Foster, & Temple, 1992; Temple, Foster & O'Donnell, 1984), and it is unlikely that the procedures used were too difficult, as the hens had learned to discriminate between two differently coloured stimuli. However, it is unclear why three of the hens learned to discriminate between the differently shaped stimuli, and why three did not learn. In order to be able to assess if hens show transfer between photographs and objects, they must first learn to discriminate between two objects (or photographs). Only when this occurs can they also show transfer of their performance to the other form of the stimuli. While half of the hens did not learn the discrimination in the present study, it is possible that they could all learn to discriminate stimuli that appear to humans to be a simpler discrimination, for example, rectangular and triangular blocks.

EXPERIMENT 8

Three of the six hens failed to learn to discriminate between the stimuli used in Experiment 7. Although the stimuli appeared to differ to humans, it seems they did not appear to differ for some hens. It is unclear why some hens learned the discrimination and some did not.

While the previous studies in this thesis have shown that hens can learn to discriminate between differently colour stimuli, they appear to have more difficulty learning to discriminate differently shaped stimuli of the same colour. They either failed to learn the discrimination (Experiment 5), learned to discriminate based on other cues such as features of the experimental equipment (Experiment 6), or only some of hens learned the discrimination (Experiment 7). It is possible that the hens in the present experiment could learn to discriminate stimuli that appear to humans to be a simpler discrimination. It was shown in Experiments 2 and 3 of this thesis that hens can learn to discriminate between simple geometric shapes (plus and circle), and so the stimuli used in this experiment were similar to the simple shapes used in Experiments 5 and 6 (rectangular shaped and triangular prism blocks).

Method

Subjects

The subjects were the same six hens used in Experiment 7.

Apparatus

The apparatus was the same as that used in Experiment 7, except the stimuli used were eight green children's playing blocks. Four of the blocks were rectangular shaped measuring 50-mm long × 38-mm wide × 38-mm high. The remaining four blocks were triangular prisms measuring 50-mm long × 42-mm wide × 38-mm high (shown in Figure 8.1). Photographs of the objects were adjusted in the same manner as Experiment 7 so as to be the same size and colour as the objects.

Figure 8.1. Images of the green blocks used throughout this experiment.

Procedure

The procedure was the same as in Experiment 7. All of the hens were trained with the objects as the target stimuli. When a triangular prism block was the stimulus, a peck to the left response key was correct, and when a rectangular block was the stimulus, a peck to the right response key was correct. The VR requirement was increased over a period of 28 sessions until all hens were responding on a VR 4 schedule of reinforcement. Hens 553, 555 and 556 completed test sessions as their accuracies were above chance. That is, there was a minimum of 100 trials in a session, therefore any percentage above 59 % was significantly different from chance at a .05 level of significance (binomial test). Over the last five training sessions before the test sessions, Hens 553's performance ranged between 58 and 76 %, of which four sessions data were significantly above chance, and therefore she also completed the test sessions.

Test sessions. Test sessions followed the same procedure used in discrimination training except the photographs were used and were presented in the same wedges used during training. The hens received three test sessions each and there were at least two training sessions between every test session.

Results

Figures 8.2a and 8.2b shows the overall percentage correct and log *d* estimates plotted over training sessions for all hens. Log *d* and percentage correct follow the same pattern. It can be seen that Hens 551, 552 and 554's accuracies all remained close to chance levels (50 %), throughout all of the training sessions. Hen 553's percentages correct were significantly above chance (above 59 %, binomial test), however, her performance was not trending upwards. Hens 555 and 556's percentages correct were at or above 85 % for five sessions after 25 and 31 sessions respectively. Performance for these hens generally remained high (around 85 %) for most of the remaining training sessions. Test sessions started after 90 training sessions for Hens 553, 555 and 556, as their percentage correct were all above 70 % (chance levels).

Figures 8.3a and 8.3b present the percentages correct to each of the stimuli and the log *c* estimates across all sessions, for all hens. Hen 556 showed a bias towards responding on the left key which was associated with the triangular prism

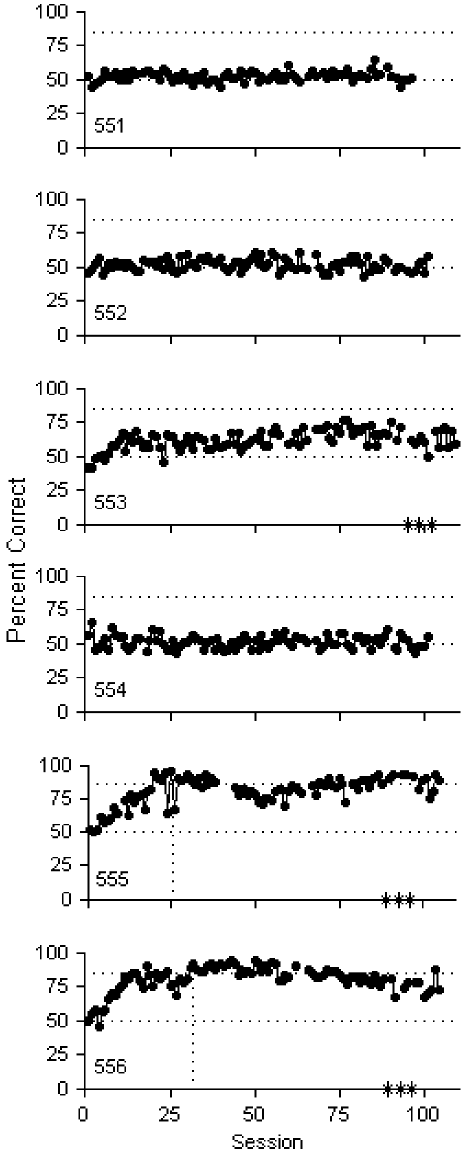


Figure 8.2a. Overall percentage correct over the training sessions with the green blocks, plotted against session number, for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal dotted lines mark 50 % and 85 %. The asterisks on the x axis represent the locations of the test sessions.

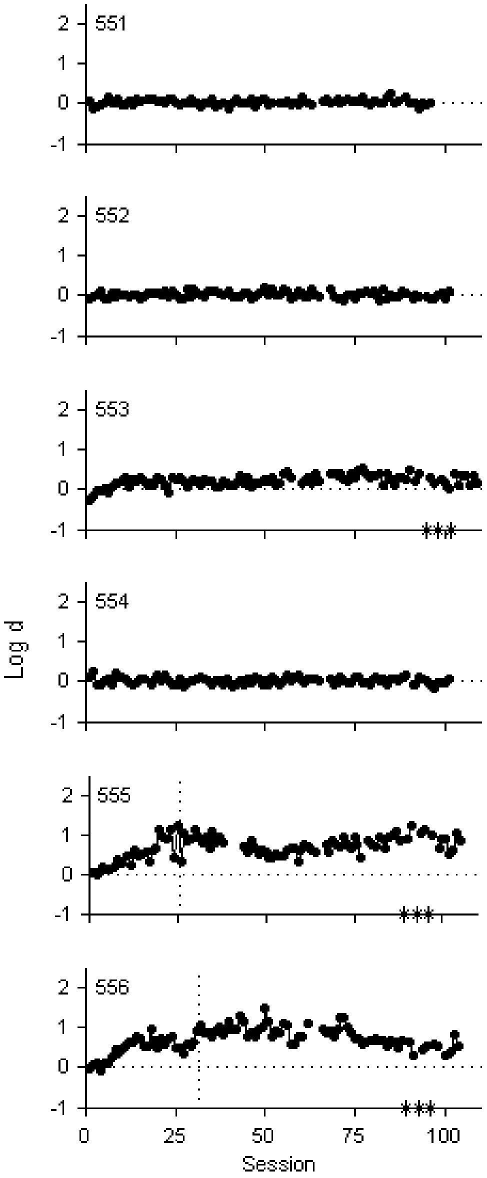


Figure 8.2b. Log *d* estimates over the training sessions with the green blocks, plotted against session number, for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

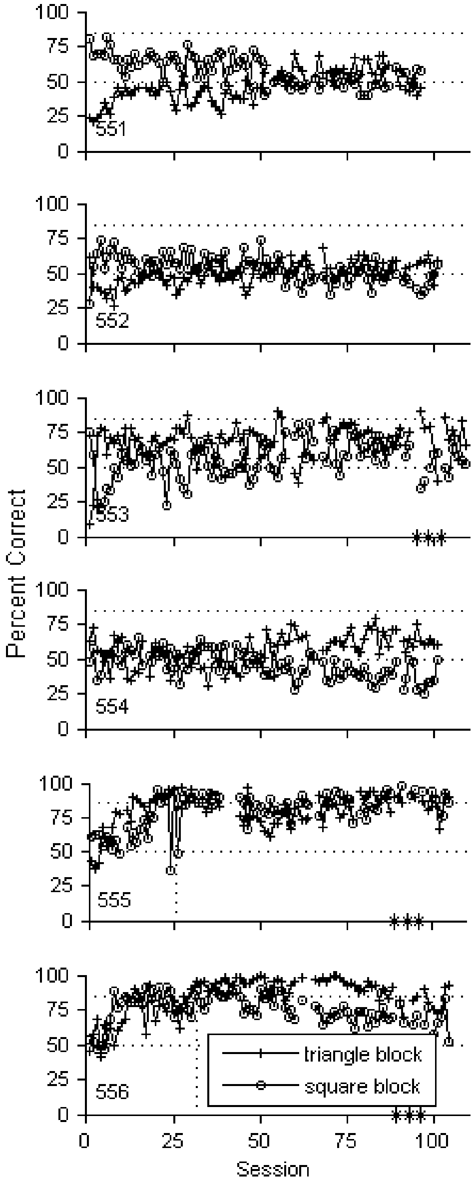


Figure 8.3a. Percentage correct towards each of the stimuli over training sessions with the green blocks for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The horizontal lines mark 50 % and 85 %. The asterisks on the x axis represent the locations of the test sessions.

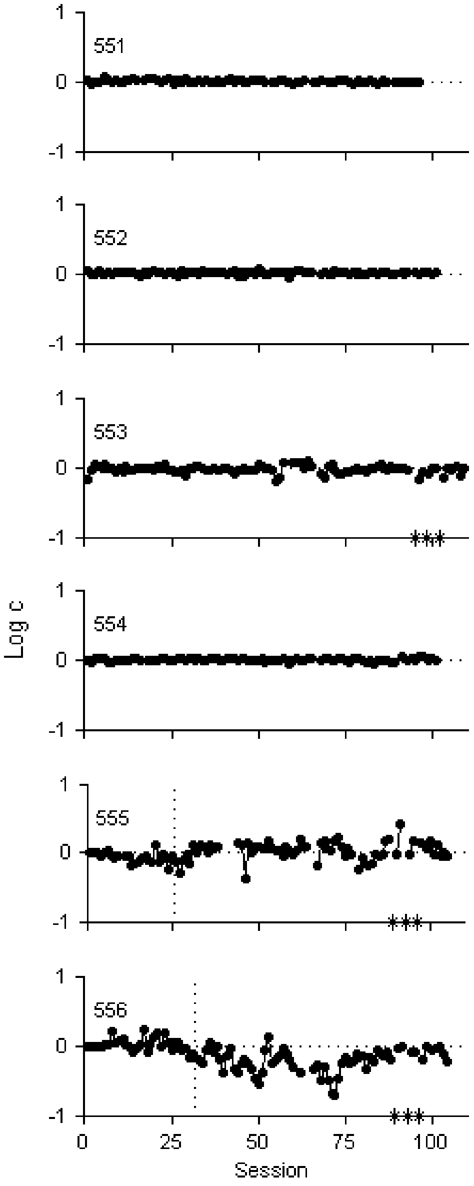


Figure 8.3b. Log *c* estimates over training sessions with the green blocks for all hens. The vertical dotted line marks the fifth session that each hen responded at or above 85 %. The asterisks on the x axis represent the locations of the test sessions.

block and was slightly more accurate with the triangular prism block (range, 82 to 92 % over the last five training sessions before test sessions) than with the rectangular block (range, 68 to 76 %). None of the other hens showed biases towards one key over the other. Hen 554 was slightly more accurate (range, 60 to 65 % over the last five training sessions) with the triangular prism block than with the rectangular block (range, 24 to 49 %) indicating an overall left key bias.

Figure 8.4 shows the mean percentage correct and standard error of the means over the last five training sessions and the data from the test sessions for Hens 553, 555 and 556. For Hens 553 and 556, percentage correct was low (range, 54 to 57 % over the six data points) over all three test sessions. Hen 555's accuracy for all three test sessions (range, 60 to 71 %), was significantly above chance (that is above 59 %, binomial test), however, these were still lower than her percentages correct during the training sessions. A repeated-measures ANOVA showed there was a statistically significant difference across the data from the test and training sessions ($F(3,6) = 13.542$, $p<.05$, $\eta^2=.871$ (large effect size)). Paired- sample *t*-tests show that there was a significant difference between the training data and data from the first and third test sessions ($t(2) = 4.687$, $p < .05$, $d=2.892$ (large effect); $t(2) = 5.907$, $p < .05$, $d=4.275$ (large effect) respectively). However, there was no significant difference between the data from the second test session and training data ($t(2) = 3.639$, $p > .05$, $d=3.175$ (large effect)). There were also no significant differences in percentage correct between any of the test sessions (test 1 and test 2, $t(2) = .988$, $p > .05$, $d=1.632$ (moderate effect); test 1 and test 3, $t(2) = 1.890$, $p > .05$, $d=1.555$ (moderate effect); test 2 and test 3, $t(2) = -.622$, $p < .05$, $d=-.741$). There was no significant trend for any of the hens (Theil test, C = 4, 2 and 2, for Hens 553, 555 and 556 respectively, $n= 4$, $p>.05$).

Discussion

The aim of this experiment was to determine if hens could learn to discriminate shapes (rectangular and triangular prism blocks) that were expected to be simpler and easier to discriminate than the Lego™ objects used in Experiment 7. However, only three of the hens showed any degree of discrimination, and only two of those hens were consistently above 85 % correct. Of these three hens, two showed no transfer of accuracy to the photographs during test sessions. Although Hen 555's accuracy was lower during the test sessions than her accuracy during training, her

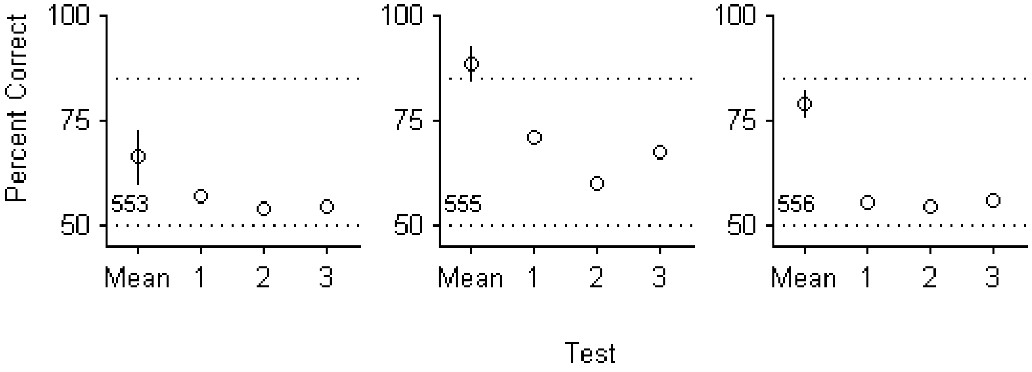


Figure 8.4. Mean percentage correct and standard error of the means over the last five training sessions with the green blocks for those hens that received test sessions. Also plotted are the data from the three test sessions (photographs). The horizontal dotted lines mark 50 and 85 %.

percentage correct remained above chance during the test sessions, showing some degree of transfer from the objects to the photographs.

As previously stated, the results found in the Experiment 7 showed that the earlier modifications to the equipment and procedure to control for extraneous variables had been successful. Thus, it seems reasonable to suggest that the hens in this experiment had also learnt the discrimination based on the stimuli that were presented to them. If these hens had learned to discriminate based on some unintended cue in the previous experiment, accuracy would have remained high despite the change in stimuli. However, all hens in the present experiment responded at chance at the start of the condition. In addition, during transfer tests, all hens' accuracies were lower than during training, indicating that these hens had not learned unintended cues in the previous experiment.

It appears as though, with these rectangular and triangular stimuli, one hen (Hen 555) did see some aspect of the photographs as equivalent to the objects. However, while she did show some degree of transfer, her accuracy did decrease during the test sessions, showing that transfer was not perfect.

The results here suggest that the current discriminative task was not in fact simpler for the hens to learn. Hens 551, 552 and 554's accuracies all remained close to chance levels (50 %), throughout all of the training sessions. This is consistent with their performances in Experiment 7 where these hens' accuracies were typically around 50 %, showing that the stimuli selected in this experiment were as difficult for the hens to discriminate as those used in Experiment 7.

These results may be best understood in conjunction with those results found in Experiments 5 to 7. For transfer of discrimination between objects and pictures to be shown, two things are required. First, the animal must learn to discriminate the stimuli presented in training. Only when this is achieved can they also show transfer of performance to the other form of the stimuli. Both of these occurred across all hens in Experiment 6, Condition 1, with the differently coloured shapes. That is, hens showed transfer of their discrimination when colour cues were also present. However, transfer from one coloured stimulus to another is not really evidence that the images were equivalent to the stimuli. These findings would be better interpreted as control by colour. In Experiments 5, 7 and 8, and Condition 3 of Experiment 6, the stimuli were such that discrimination was possible only on the shape of the object. None of the hens learned to discriminate the objects in Experiment 5. All learned to discriminate the stimuli to some degree in Condition 3 of Experiment 6, however, further testing showed that the hens' behaviour was under the control of extraneous features of the experimental equipment, rather than the stimuli. Only half of the hens learned the discrimination to any degree in Experiments 7 and 8. For those hens that did learn the discrimination, none transferred that discrimination from the objects to the photographs (Experiment 7, Condition 1 and Experiment 8), or from the photographs to the objects when they had been trained with the photographs (Experiment 7, Condition 2). It appears that the photographs did not act as substitutes for the objects (and vice versa) for these hens when shape was the only means with which to distinguish between them.

It is possible that the three non-discriminating hens in these experiments might have shown higher accuracy with even more extended training. However, this possibility seems unlikely given that the two hens (555 and 556) that learned the discrimination in Experiment 8 to above 85 % did so reasonably quickly (within 31 training sessions), and those hens that learned the task (i.e., were responding above chance) in the three conditions of Experiment 7, did so within 20-62 training sessions. In both of these experiments, the possibility that the hens had learnt to discriminate based on some other cue could be ruled out as the equipment and procedure modifications appear to have controlled for extraneous cues that had been learned in earlier experiments. If the hens were using some feature of the equipment, other than the stimuli, in their discrimination, it would be expected that they would retain that discrimination during test sessions, as the same wedges were used to present the stimuli during training and testing. However, all hens (except Hen 555) in this experiment responded at around chance during the test sessions, indicating that they were indeed discriminating in the training sessions based on the stimuli. The three non-discriminating hens in the present experiment had approximately 100 training sessions and their performances remained close to 50 % over all sessions, indicating that further training sessions would probably not have lead to improved accuracy.

The photograph stimuli used in all experiments of this thesis were controlled so that they were as close as could be in colour and size as to the real object. Spetch and Friedman (2006) also controlled for size and colour of the stimuli and presented them from different views. They found some degree of transfer to the alternative stimuli and possible evidence of learning in that the proportion correct increased over trials. However, Dittrich et al., (2010) point out that the pigeons' accuracy during initial transfer tests was lower, indicating that transfer was not perfect. Transfer was shown by only one hen in the present study, and this was lower than accuracy shown during training. The rest of the hens showed no transfer of discrimination from objects to photographs, and vice versa. It is not clear what could result in the differences in findings by this study and those of Spetch and Friedman, except that the different stimuli used may have lead to the different findings. In addition, pigeons were used in the study by Spetch and Friedman, and hens were used in the present study. It may be that pigeons do see images as equivalent to objects, whereas hens do not.

That the hens in this study did not transfer a discrimination to pictures of objects is in contrast with those of Cabe (1976), Lumsden (1977), and Spetch and Friedman (2006), who all found pigeons showed transfer of discrimination between objects and their pictures. However, there are also a number of studies that found hens did not transfer a discrimination from objects to their images (e.g., Bradshaw & Dawkins, 1993; Candland, 1969; Weavers, 2000) which is in line with the results of this thesis. It is interesting to note that those studies that have found successful transfer all used pigeons as subjects, and those studies failing to find evidence of transfer used hens. It seems as though pigeons may be able to transfer a discrimination to photographs, whereas hens do not. Thus, while it seems unlikely, it may be possible that hens and pigeons perceive pictorial images differently. Pigeons have been shown to discriminate between stimuli (e.g., presence of humans in pictures, Herrnstein & Loveland, 1964; presence of fish in underwater pictures, Herrnstein & de Villiers, 1980; different types of leaves, Cerella, 1979; pictures of cats and dogs, Ghosh, Lea & Noury, 2004; and cubist and impressionist paintings, Watanabe, Sakamoto & Wakita, 1995). Clearly, pigeons seem able to perform quite complex discriminative tasks.

As previously mentioned, Premack (1976) proposed that for images to represent objects for an animal, the animal must respond to both images and objects in the same way, but also respond to each differently. However, images are typically used in animal research, not to be a representation of a stimulus in the way Premack suggests, but as a substitute for that stimulus. Thus, when an image is to be used as a substitute for the real stimulus, the animal must respond as if the images and real stimuli are equivalent, not different. If an animal responds to both objects and images in a similar manner, this shows confusion between the object and the picture as the animal responds as if the image is the object. In this thesis, the results from Experiments 5, 7 and 8 show that while three of the hens learned to discriminate between the stimuli (both objects and photographs), they did not transfer that discrimination to the alternative stimulus (photograph or object). This finding shows that the hens showed neither confusion nor correspondence between the objects and their images. Rather, it appears that these hens treated the images and the objects as independent of each other. Fagot et al. (1999) outlines independence as occurring when the animal shows no association between the objects and their pictures, and states that this shows that ‘processing’ of the pictorial stimuli is done independently of that of the real object. If the hens were treating objects and pictures as independent, then it is not surprising that transfer of discriminative performance did not occur in the present study.

SUMMARY

The first part of this thesis aimed to establish whether it could be that the types of screens that have often been used to present images to animals affect hens' ability to recognise these images. It was argued initially that whether or not the use of CRT screens might affect the image an animal sees could be related to the animals' critical flicker fusion (CFF) frequency. Thus, in Experiment 1, the CFF frequency for hens at one particular luminance level (300 cd/m^2) was established using two stimulus discrimination presentation methods (successive and simultaneous). The CFF values across the two procedures gave comparable results showing that either stimulus presentation method can be used to assess CFF. In addition, it was found that hens' CFF values ranged from 68.5 to 95.4 Hz, which are higher flicker fusion values than those of humans at that particular luminance. This finding was in line with those found by Jarvis et al. (2002) and Nuboer et al. (1992). It suggests that hens (and any other animal with high CFF values) may perceive stimuli presented to them on computer or television monitors that have refresh rates lower than their CFF, particularly CRT monitors, as flickering. It is possible that TFT monitors may be a more appropriate method to present stimuli as they are virtually flicker free.

There has been a recent move in research towards using TFT monitors because they do not flicker as CRT monitors do, and because they have become more cheaply and readily available. Experiments 2 and 3 of this thesis assessed whether hens could transfer a discrimination learned on a TFT monitor to a CRT monitor set at different refresh rates. In Experiment 2, the stimuli were different colours and different 2D shapes. The hens quickly learned to distinguish between two colours (red and green) presented on a TFT monitor and transferred this discrimination to a CRT monitor set at 60 Hz. When the stimuli were two shapes (circle and cross), stimulus control was not attained until an FR 5 observing response was included. During transfer tests to these same stimuli, but now presented on a CRT monitor set at a range of refresh rates, all hens' accuracy decreased (especially at 60 Hz) suggesting that the stimuli did not appear the same as the training stimuli, although there was evidence of some transfer. The testing procedure was not the same as the training procedure as the observing response could not be recorded using the apparatus that was available.

In Experiment 3, the apparatus was altered by attaching infrared screens around the front of the TFT and CRT monitors that allowed the observing response to be recorded. This allowed the testing and training procedures to be identical. In Experiment 3, the stimuli were different colours, 2D shapes, pictures of Lego^{TM} blocks, and line drawings. The hens did not learn to discriminate pictures of the Lego^{TM} blocks in Experiment 3, and so transfer could not be measured with these stimuli. Also, only half of the hens learned to discriminate either the real Lego^{TM} blocks or the photographs of the blocks used in Experiment 3 in Experiment 7. Thus, it appears that these particular stimuli were difficult for the hens to discriminate, regardless of the manner of stimulus presentation.

In spite of the failure with the Lego^{TM} stimuli, the hens did learn to discriminate the different shapes (plus and circle), colours (blue and yellow), and line drawings (watering can and iron) presented on a TFT monitor. Transfer tests to the CRT monitor with these stimuli at different refresh rates showed that hens were generally highly accurate with refresh rates above their CFF, but were not so accurate at the lower refresh rates. Presumably, the stimuli appeared increasingly different from the training stimuli as the refresh rate decreased. As previously mentioned, there are at least two possible reasons for this decrement in accuracy at the lower refresh rates in Experiments 2 and 3. Firstly, the images may have been difficult for the hens to see at the low refresh rates. That is, the discrimination was the same, but made more difficult by the change in flicker, and the hens' visual systems may be limited in ability to discriminate stimuli close to or below their CFF. Secondly, the images may not be immediately recognisable as the training stimuli. That is, the flickering stimuli may appear to be different from the training stimuli making the discrimination task during the test sessions a new one. To test this, in Experiment 4, the hens were trained to discriminate different line drawings (boat and lamp) presented on a CRT monitor set at 60 Hz (which was below the hens' CFF and assumed to appear as flickering). Transfer tests consisted of presenting the same stimuli on the CRT monitor set at higher refresh rates, and on the TFT monitor. It was found that accuracy remained high (above 85%) across all test sessions with the CRT monitor, and was relatively high (around 80%) with the TFT monitor. The discrimination learned at 60 Hz then transferred to steady images. Hence, although hens have difficulty transferring a discrimination from steady to flickering stimuli, they can be trained to discriminate flickering stimuli, and this implies the hens are able to see the flickering stimuli. The results from Experiments 2 to 4 suggest that transfer of a discrimination from a flickering to a steady stimulus occurs more readily than transfer from a steady to a flickering stimulus. Therefore, the screen type and refresh rate used in training both have an effect on the degree to which hens transfer a discrimination.

The second part of this thesis aimed to assess if hens responded to photographs in the same way that they did to the real objects that were depicted in the photographs, and vice versa. Typically, studies assessing correspondence between objects and pictures measure natural spontaneous responses to biologically relevant stimuli (e.g., courtship behaviour). There are relatively few studies that examine correspondence by measuring transfer of a learned response across stimuli. Experiment 5 established a procedure to investigate whether hens transferred a discrimination between two 3D objects to 2D photographs of the same objects, and vice versa. Hens 51 to 53 were trained with the objects, and Hens 54 to 56 were trained with the photographs. None of the hens showed high discriminative performance, and all showed strong biases in their responding. The failure to learn to discriminate may have resulted from the stimuli appearing to be too similar to the hens.

As such, in Experiment 6, a colour/shape discrimination was used (with both objects and photographs), and all of the hens learned the discrimination quickly, showing high accuracy and no biases with the different coloured stimuli. Experiment 5 raised another issue. It was possible that the stimuli used there may have been presented too far away for the hens to see the stimuli clearly, and so the stimuli were moved closer in the second condition in Experiment 6. This change in viewing distance resulted in an initial disruption to the hens' accuracies, but most hens quickly relearned the discrimination. All hens showed transfer of their discriminative performance from photographs to objects and from objects to photographs at both viewing distances. In the last condition in Experiment 6, the hens appeared to learn to discriminate between the stimuli (objects or photographs) that were different shapes, but the same colour. However, a series of test sessions found that, for most hens, behaviour had come under control of unintended features of the experimental equipment other than the stimuli. Thus, the equipment and procedure were modified to control for these extraneous cues in the following experiments.

The stimuli used in Experiment 7 were two different pairs of Lego^{TM} blocks of the same colour. In Condition 1, the stimuli were the real blocks, and in Condition 2, the stimuli were the photographs of the pairs of blocks. In both conditions, only three of the six hens learned the discrimination to any degree, while the remaining three hens' accuracies remained close to chance. Of those hens that showed some discrimination, none transferred their accuracy to the alternative stimuli (photographs in Condition 1, and objects in Condition 2) indicating that the photographs were not seen as equivalent to the real objects, and vice versa. In Condition 3, the real objects that had been presented as photographs in Condition 2 were used to see if the hens learned the discrimination more quickly if they had previous experience with the photographs. However, there was no statistical difference in the number of sessions required to respond consistently above 85 % across Conditions 2 and 3, showing that the hens did not learn to discriminate the objects more quickly if they had previous experience learning to discriminate the photographs. This finding also implied that the photographs and objects were not seen as equivalent. The modifications to equipment and procedure to prevent behaviour coming under control of extraneous cues of the wedges used to present the stimuli appeared to have been successful. This was shown in the hens decrease in accuracy during the transfer tests. If the hens had learned to discriminate using extraneous cues, as they had in Condition 3 of Experiment 6, accuracy would have remained high during the test sessions, as both training and testing sessions used the same wedge sections of the wheel. However, accuracy decreased for all hens, showing that the hens must have been discriminating based on the shapes.

As only three of the hens learned to discriminate the shapes used in Experiment 7, it appears that these were difficult discriminations to learn. In an attempt to get all six hens to learn the discrimination, stimuli that appeared to be a simpler discrimination to humans were used in Experiment 8. Again, the same three hens learned the discrimination to some degree as in Experiment 7, and the same three hens failed to learn showing that the stimuli used in Experiment 8 were not easier to discriminate. Only one of the hens showed some degree of transfer to the photographs, although accuracy during the test sessions was much lower than accuracy shown during training. It is possible that this hen had learned to discriminate based on some unintended cue that also transferred during the test sessions, rather than discriminating the stimuli that were presented. However, this seems unlikely as accuracy for this hen was lower during the test sessions than during training, and her accuracy decreased to around 50% with each new stimulus presentation in Experiments 7 and 8. If she had learned to discriminate based on some other feature, accuracy would have remained high across the new stimuli and across transfer tests. In addition, none of the other hens learned extraneous cues after the equipment had been modified.

GENERAL DISCUSSION

One issue with the methods used to present visual stimuli to animals is that they have been designed for the human visual system. Researchers have to be careful not to ignore the fact that animals can have very different visual systems. Thus, if the images are to be used as a substitute for the real objects, it is necessary to determine if, and how, animals see pictures, and establish what factors may affect their ability to see pictures. This leads to difficulties in the interpretation of research that has used pictures in place of real stimuli without having first examined whether animals see them as the same (Fagot et al., 1999).

One implication of the finding that hens have high CFF thresholds is that images that have been presented on CRT monitors set at low refresh rates may be difficult for hens, and any animals with high CFFs, to see. Thus, TFT monitors may be a more appropriate means with which to present stimuli to animals that have high CFFs. While hens learned to discriminate both steady and flickering images, transfer from a steady to a flickering stimulus does not appear to occur as readily as transfer from a flickering to a steady stimulus. The data from Experiments 2 to 4 suggest that the use of CRT monitors at low refresh rates in some research may affect accuracy in a discrimination task. This is the only experiment, to the author's knowledge, to assess how refresh rate affects discrimination and to directly compare discriminative performance across two different types of screens.

Transfer tests often involve presenting the alternative stimuli for a percentage of trials throughout the normal training procedure. In this thesis, transfer tests consisted of a whole session in which correct responding was still reinforced. Typically, only one transfer-test session was used in Experiments 2 to 4, assessing transfer across monitors and refresh rate. This was done to reduce the chances that the hens would learn the new discrimination, as reinforcement was still available during test sessions. It was considered that any transfer, or lack thereof, would be shown immediately once the situation was changed and did not require more than one session. It was also considered that having more than one test session might result in leaning of the new discrimination. Examination of the within-session data showed that learning did not occur during the test sessions, therefore more test sessions could have been conducted. However, as there were at least 100 trials during test sessions, more transfer-test sessions would not have added anything further to the findings. In Experiments 5 to 8, hens completed up to four test sessions, but examination of the data suggests that more than one was not necessary. If accuracy had increased over test sessions, this would more likely have been an indication that hens were learning the new discrimination rather than that they were showing true transfer from one stimulus to the other. Regardless, accuracy did not increase over these later test sessions (except possibly for two hens in Condition 1 of Experiment 7) showing that learning did not readily occur in these studies as a result of these extra test sessions.

In studies assessing picture-object correspondence in animals, findings have often been equivocal and contradictory, particularly with birds. It may be that some birds are able to show correspondence between objects and pictures (see Cabe, 1976; Lumsden, 1977; Spetch & Friedman, 2006). However, many studies that have claimed to show correspondence failed to control other factors that may have aided the discrimination (e.g., size differences (Watanabe, 1993), texture cues (Watanabe, 1997), and picture orientation (Looney & Cohen, 1974)). The hens in the present study transferred their discriminative performance from objects to photographs (and vice versa) when colour cues were available. However, they did not transfer their discrimination when it was based, as best as could be arranged, on shape alone. These findings indicate that pictures might possibly be used as visual stimuli when discrimination is based on cues such as colour. However, pictures may have a limited use if they are to be used a substitutes for real stimuli. It remains to be seen if transfer occurs when there are many differences between stimuli (e.g., size and texture). However, there are a number of studies that failed to find transfer when using stimuli that could be discriminated on a number of features, such as conspecifics or food (e.g., Bradshaw & Dawkins, 1993; Dittrich, et al., 2010; Trillmich, 1976; Watanabe et al., 1993; Weavers, 2000). As previously stated, researchers cannot assume that animals are seeing visual stimuli in the same way that humans do and that picture-object correspondence automatically occurs. Transfer of discrimination between pictures and objects does not appear to be automatic or simple. In fact, it is quite possible that animals need to be trained to do it.

Even in humans, recognition of pictorial stimuli requires prior experience with such stimuli. Cross-cultural studies have shown that members of cultures that have had little exposure to pictures may initially have difficulty in recognising the images portrayed in the pictures as the objects, however, they could learn to do so when the features of the pictures were pointed out (Deręgowski, 2000; Miller, 1973), yet they do not confuse the pictures for the real objects. Truppa et al. (2009) note that the extent of an animal's picture-object correspondence may only be evident after receiving some form of pre-training with the pictures. Thus, expecting animals to respond automatically to pictures as they would to real world objects or animals is problematic. But it might be possible to train animals to associate objects with their images. It is known that hens can do matching-to-sample tasks (e.g., Foster et al., 1995; Nakagawa et al., 2004; Weavers, Foster & Temple, 1998). Therefore, a possible procedure to do this may be to train animals in a matching-to-sample task where the sample stimuli are either 2D pictures or 3D objects, and the comparison stimuli are the opposite stimuli. In this way, animals could be trained to associate objects with their pictures, however, this does not mean the animals will necessarily see the pictures as substitutes for the objects. While it may be possible to train animals to associate pictures with the real objects, the question arises whether the animal sees the pictures as a substitute for, or a representation, of the real objects. Most research that present pictures as stimuli actually wish to examine the animals' responses to the real objects and therefore want the animals to respond to the pictures in the same way as they do to the objects. That is, to confuse the two stimuli, so that the pictures can be used as a substitute. If animals do see the images as a representation, then they will respond to them differently from the way they would to the real stimulus (as suggested by Premack, 1976), and so the image would not produce the same behaviour as the real stimulus. That is, while humans see pictures as representations of the real world, they do not confuse them and respond to them as if they are real. For example, humans can name pictures of food items and may even salivate, yet they do not try and eat them.

A problem with research in visual perception, is that there is a possibility that other, unintended, discriminative cues may correlate with the stimuli used, and thus come to control behaviour, instead of the intended stimulus. The results from Condition 3 of Experiment 6 show that, even with stimuli that appeared quite different to humans, the hens learned to discriminate based on what appeared to be quite 'subtle' cues. It may be that these cues were more 'obvious' to the hens. Thus, care needs to be taken that there are not other cues, other than the one intended, that animals may be responding to. This could be tested by removing the stimuli, and seeing if accuracy remains high, or reduces to 50 %.

It must be noted that there are difficulties when selecting stimuli for this type of research. The stimuli used throughout the last four experiments in this thesis were specifically selected to try and ensure the discrimination could be based on shape alone (except for those when colour was a varied dimension). This was done in an attempt to make the shape of the object itself the subject of the discrimination rather than some other property of the object such a texture, colour or size. In addition, 3D cues were eliminated by lighting the stimuli in such a way that shadows were not present. The shapes used here were more similar to each other than those used by Spetch and Friedman (2006). The stimuli used by Spetch and Friedman contained many component shapes (see Figure 5.1) and also contained 3D cues such as shadows. They were also presented from a number of different viewpoints. This meant that the pigeons were required to attend to more than one feature of the stimuli. In the present study, the hens were only presented with one viewpoint of each stimulus and shadows were eliminated. It is possible that the hens used in this thesis may have learned to discriminate stimuli more readily, and may have shown some transfer, if the shadow cues had been kept, and /or if different viewpoints had been used. Thus, more research is required on how different viewpoints, or the provision of shadows, may affect hens' ability to learn a discrimination, and to transfer that discrimination to alternative stimuli.

As the Lego^{TM} stimuli were difficult for hens to discriminate when presented on computer monitors (Experiment 3), and when shown as photos or objects (Experiment 7), this suggests that the lack of transfer could have been a result of the stimuli, and not of an inability of the hens to see the photographs as a substitute for the objects. Although these stimuli appeared different to the human eye, it seems they may not appear different to hens. It is unlikely that the hens were unable to see the stimuli clearly, as DeMello et al. (1992) showed that hens have high visual acuity. Therefore, it is not obvious why the hens were unable to learn this particular discrimination. When stimuli differ on more than one feature (e.g., colour, size, texture), transfer can occur on any (or all) of these features. For example, if a hen was trained to discriminate a grain of wheat from a grain of rice there are a number of features that the hen could use to discriminate between the two foods. If transfer to pictures was shown, this could be based on any of these features, and further testing would be required to discover which and to rule any out. However, this would not mean that the animals were not seeing some aspect of the pictures as equivalent to the real stimulus. Had stimuli that differed in more than one aspect been used in the last experiments of this thesis, the results might have been different. That is, hens may have shown transfer of performance to photographs, and vice versa. However, it was not clear what stimuli should have been used. Given the difficulty in finding stimuli that all hens could discriminate between, further research is required with a range of stimuli to find stimuli that all hens can distinguish. It may not be that hens are unable to transfer a discrimination, but it may be that whether or not they can depends on the stimuli that are used. It would be interesting to examine transfer using stimuli that hens have previously been shown to discriminate. For example, Experiment 3 showed that hens readily discriminated between a cross and a circle and a pair of line drawings. It is possible that had 3D versions of these stimuli been used hens may have shown transfer to the 2D images. The procedure developed here could be used to test for transfer of a discrimination using a range of stimuli.

In light of the finding that including an observing response aided in hens' learning to discriminate between stimuli in Experiment 2, it is possible that increasing the time required for the observing response (the 0.5 s beam break) may have aided discriminative accuracy in Experiments 5 to 8. While this would have increased the length of each trial, and thus decreased the number of trials per session, any decrease in reinforcement rate could have been offset by the increased reinforcement rate gained through a possible increase in accuracy. As previously stated, DeMello, Foster and Temple (1993) and White (1985) found that increasing an observing response requirement from FR 1 to FR 5 increased accuracy on a visual acuity task with hens. However, DeMello et al. also found that further increases in this response requirement greater than FR 10 did not further increase accuracy for all hens. That is, while for some hens, accuracy increased further, for others accuracy was either maintained or decreased. The effect of the length of observing responses on discriminative performance in the present procedure merits further study.

In this thesis, static images were presented on computer screens in Experiments 2 to 4, and photographs of objects, and the objects, were presented in Experiments 5 to 8. Although, at first glance, static pictures may bear no relation to moving images, they have often been used in a similar manner to examine processes such as spatial memory learning (e.g., Spetch & Wilkie, 1994), social facilitation (e.g., Keeling & Hurnik, 1993), motion perception (e.g., Lea & Dittrich, 1999) and social behaviour (i.e., responses to conspecifics, e.g., Clark & Uetz, 1990; Evans &

Marler, 1991). In studies that have used photographs or video playback, the stimuli have often been used as a substitute for a real object or animal. Thus, the finding that hens did not respond to pictures that depicted real objects in the same way they did to the objects has implications for research that present moving stimuli.

Static images were used in this research, but studies that examine animals' responses to moving images are also required. Factors, such as movement cues, can be important in an animal's recognition of stimuli such as conspecifics. However, moving images, like static images, are often used without first assessing if they function as a substitute for the real object.

Often studies use video playback to present moving images of conspecifics or other species and measure an animal's response to the video. One advantage of video playback over the use of real animals is that the same sequence of behaviours can be presented to a number of subjects reducing variability that can occur if live animals are shown. However, animals cannot interact with the images using video playback, and as a result, animals may stop responding, or behave differently when no feedback from the video image is available. One way to measure birds' responses to images further would be to examine how they respond to images of live animals shown in real-time on video using two way cameras so that the animals can view each other in real time and interact. To the author's knowledge there are no studies that directly assess if birds respond to live video images of other conspecifics. If the birds respond in the same way to the live real-time images presented on a TFT monitor as they do to the real conspecifics, this would provide evidence that the image is equivalent to the real stimulus. This would be a great advantage for research with conspecific recognition, as while it is difficult to present and remove real conspecifics multiple times in a discrimination task, this could be done easily with images of the conspecifics.

Baldauf, Kullman and Bakker (2008) state that many studies do not take into account, or may be unaware of, the technical limitations of their equipment, and that often technical details of experimental equipment are not reported. As a result, there may be methodological problems present in the experimental design that the researcher is unaware of. For example, as the present research shows, CRT monitors, which have been commonly used in research, may have distorted images for animals with high CFF thresholds.

In addition, Baldauf et al. (2008) also point out that while 3D depth cues can be added to pictorial stimuli, they are still being presented on 2D devices. Therefore, whether an animal can perceive them as 3D cues may depend on how an animal perceives depth. However, it is not clear that, if an animal does not 'confuse' or substitute the images with the objects, whether the addition of 3D cues to the 2D images would aid discrimination.

In conclusion, when researchers use pictures as substitutes for real objects, they need to be careful to take into account the properties of the visual systems of the animals when selecting how to present stimuli. That is, they need to establish that the animals can perceive the stimuli that are being presented, and that the method of stimulus presentation is species appropriate. This was done in this thesis by establishing the hens' CFF and determining if the refresh rates of a CRT monitor affected the hens' discrimination of stimuli. In addition, if photographs are to be used as a substitute for real objects, then it is necessary to determine if animals respond to the photographs in the same way they do to the real objects that are depicted in the photographs. This was tested in the present experiment, and it was found that, for those hens that learned the initial discrimination, none showed transfer to either the photographs or the objects. As already pointed out, it seems as though hens do not respond to pictures as if they were the real stimuli, at least under the conditions used in this thesis and using the present stimuli. If researchers wish to use pictures as a substitute for real objects, then they should first demonstrate that the animal responds in a similar way to both the real objects and the pictures of the objects.

REFERENCES

Abeyesinghe, S.M., McLeman, M.A., Owen, R.C., McMahon, C.E., & Wathes, C.M. (2009). Investigating social discrimination of group members by laying hens. *Behavioural Processes, 81(1),* 1-13

Adret, P. (1997). Discrimination of video images by zebra finches (Taeniopygia guttata): Direct evident from song performance. *Journal of Comparative Psychology, 111(2),* 115-125.

Alsop, B., & Jones, B.M. (2008). Reinforcer control by comparison-stimulus color and location in a delayed matching-to-sample task. *Journal of the Experimental Analysis of Behavior, 89(3),* 311-331.

Aron, A., Aron, E. N., & Coups, E. J. (2009). *Statistics for psychology.* Upper Saddle River, NJ: Prentice Hall.

Autrum, H. (1950). Die Belichtungspotentiale und das Sehen der Insekten (Untersuchungen an *Calliphora* und *Disippus). Zeitschrift für vergleichende, 32,* 116-227.

Bauldauf, S.A., Kullmann, H., & Bakker, T.C.M. (2008). Technical restrictions of computer-manipulated visual stimuli and display units for studying animal behaviour. *Ethology, 114,* 737-751.

Bird, C.D., & Emery, N.J. (2008). Using video playback to investigate the social preferences of rooks, Corvus frugilegus. *Animal Behaviour, 76(3),* 679-687.

Bovet, D., & Vauclair, J. (2000). Picture recognition in animals and humans. *Behavioural Brain Research, 109*, 143-165.

Bradshaw, R.H., & Dawkins, M.S. (1993). Slides of conspecifics as representatives of real animals in laying hens (Gallus domesticus). *Behavioural Processes, 28,* 165-172.

Brodbeck, D. R., & Shettleworth, S.J. (1995). Matching location and color of a compound stimulus: comparison of a food-storing and a nonstoring bird species. *Journal of Experimental Psychology: Animal Behavior Processes, 21*(1), 64-77.

Brundett, G.W. (1974). Human sensitivity to flicker. *Lighting Research and Technology, 6,* 127-143.

Bushnell, P.J. (1999). Detection of visual signals by rats: Effects of signal intensity, event rate, and task type. *Behavioural Processes, 46(2),* 141-150.

Cabe, P.A. (1976). Transfer of discrimination from solid objects to pictures by pigeons: A test of theoretical models of pictorial perception. *Perception and Psychophysics, 19(6),* 545-550.

Candland, D. K. (1969). Discriminability of facial regions used by the domestic chicken in maintaining the social dominance order. *Journal of Comparative and Physiological Psychology, 59*(2), 281-285.

Castro, L., & Wasserman, E.A. (2010). Effects of stimulus size and spatial organization on pigeons' conditional same-different discrimination. *Behavioural Processes, 83(2),* 162-171.

Cerella, J. (1979). Visual classes and natural categories in the pigeon. *Journal of Experimental Psychology: Human perception and performance, 5*(1), 68-77.

Chen, M.T., & Lin, C.C. (2004). Comparison of TFT-LCD and CRT on visual recognition and subjective preference. *International Journal of Industrial Ergonomics, 34,* 167-174.

Clark, D. L., & Stephenson, K.R. (1999). Response to video and computer-animated images by the tiger barb, Puntius tetrazona. *Environmental Biology of Fishes, 56*, 317-324.

Clark, D. L., & Uetz, G. W. (1990). Video image recognition by the jumping spider, Maevia inclemens (Araneae: Salticidae). *Animal Behaviour, 40*, 884-890.

Cohen, J. (1994). The earth is round (*p*<.05). *American Psychologist, 49(12),* 997-1003.

Davison, M., & McCarthy, D. (1989). Effects of relative reinforcer frequency on complex color detection. *Journal of the Experimental Analysis of Behavior, 51(3),* 291-315.

Davison, M., & Tustin, R.D. (1978). The relation between the generalized matching law and signal-detection theory. *Journal of the Experimental Analysis of Behavior, 29(2),* 331-336.

Dawkins, M.S. (2002). What are birds looking at? Head movements and eye use in chickens. *Animal Behaviour, 63(5),* 991-998.

Dawkins, M.S., & Woodington, A. (1997). Distance and the presentation of visual stimuli to birds. *Animal Behaviour, 54,* 1019-1025.

D'Eath, R. B. (1998). Can video images imitate real stimuli in animal behaviour experiments? *Biological Review, 73*, 267-292.

D'Eath, R. B., & Dawkins, M. S. (1996). Laying hens do not discriminate between video images of conspecifics. *Animal Behaviour, 52,* 903-912.

D'Eath, R.B., & Keeling, L.J. (2003). Social discrimination and aggression by laying hens in large groups: From peck orders to social tolerance. *Applied Animal Behaviour Science, 84(3),* 197-212.

Delius, J. D. (1992). Categorical discrimination of objects and pictures by pigeons. *Animal Learning and Behavior, 20*(3), 301-311.

Delius, J.D., Emmerton, J., Hörster, W., Jäger, R., & Ostherim, J. (2000). Picture-object recognition in pigeons. In J. Fagot (Ed.) *Picture Perception in Animals.* Psychology Press: East Sussex, pp. 1-36.

DeMello, L.R. (1989). *The effect of procedural variations on discriminative performance: Visual acuity in the domestic hen.* Unpublished Doctorate, University of Waikato, Hamilton.

DeMello, L. R., Foster, T. M., & Temple, W. (1992). Discriminative performance of the domestic hen in a visual acuity task. *Journal of the Experimental Analysis of Behavior, 58*, 147-157.

DeMello, L.R., Foster, T.M., & Temple, W. (1993). The effect of increased response requirements on discriminative performance of the domestic hen in a visual acuity task. *Journal of the Experimental Analysis of Behavior, 60(3),* 595-609.

Deręgowski, J.B. (2000). Pictorial perception: individual and group differences within the human species. In J. Fagot (Ed.) *Picture perception in animals*. Psychology Press: East Sussex, pp. 397-429.

DeVoe, R.D. (1963). Linear relations between stimulus amplitudes and amplitudes of retinal action potentials from the eye of the wolf spider. *The Journal of General Physiology, 47*(17), 13-32.

Dittrich, L., Adam, R., Ünver, E., & Güntürkün, O. (2010). Pigeons identify individual humans but show no sign of recognizing them in photographs. *Behavioural Processes, 83*, 82-89.

Emmerton, J., & Renner, J.C. (2006). Scalar effects in the visual discrimination of numerosity by pigeons. *Learning & Behavior, 34(2),* 176-192.

Evans, C. S., & Marler, P. (1991). On the use of video images as social stimuli in birds: audience effects on alarm calling. *Animal Behaviour, 41*, 17-26.

Fagot, J., Martin-Malivel, J., & Dépy, D. (1999). What is the evidence for an equivalence between objects and pictures in birds and nonhuman primates? *Current Psychology of Cognition, 18,* 923-949.

Ferguson, C.J. (2009). An effect size primer: A guide for clinicians and researcher. *Professional Psychology: Research and Practice, 40(5),* 532-538.

Fleishman, L. J., Marshal, C. J., & Hertz, P. E. (1995). Comparative study of temporal response properties of the visual system of three species of anoline lizards. *Copeia, 2,* 422-431.

Foster, T.M., Temple, W., MacKenzie, C., DeMello, L.R., & Poling, A. (1995). Delayed matching-to-sample performance of hens: Effects of sample duration and response requirements during the sample. *Journal of the Experimental Analysis of Behavior, 64(1),* 19-31.

Friedrich, A.M., & Zentall, T.R. (2010). A relational differential outcomes effect: Pigeons can classify outcomes as “good” and “better”. *Animal Cognition, 13(2),* 359-365.

Galoch, Z., & Bischof, H. (2006). Zebra finches actively choose between live images of conspecifics. *Ornithological Science, 5*, 57-64.

Galoch, Z., & Bischof, H. (2007). Behavioural responses to video playbacks by zebra finch males. *Behavioural Processes, 74*, 21-26.

Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour, 66,* 15-36.

Ghosh, N., Lea, S.E.G., & Noury, M. (2004). Transfer to intermediate forms following concept discrimination by pigeons: chimeras and morphs. *Journal of the Experimental Analysis of Behavior, 82(2),* 125-141.

Ginsberg, N., & Nilsson, B. (1971). Measuring flicker thresholds in the budgerigar. *Journal of the Experimental Analysis of Behavior, 15,* 189-192.

Greene, S. L. (1984). Feature memorization in pigeon concept formation. In *Quantitative analyses of behavior: Acquisition processes* (eds. M. Commons, A. R. Wagner & R. J. Herrnstein), Vol 4, pp. 209-229. Cambridge, MA: Ballinger.

Guttman, N., & Kalish, H.I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology, 51(1),* 79-88.

Hart, W.M. (1992). *The temporal responsiveness of vision.* In: F.H. Adler & W.M Hart (Ed.), Adler’s Physiology of the Eye: Clinical Application, 9th edn. Mosby, St. Louis.

Hauser, J., & Huber-Eicher, B. (2004). Do domestic hens discriminate between familiar and unfamiliar conspecifics in the absence of visual cues? *Applied Animal Behaviour Science, 85,* 65-76.

Hautus, M.J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of *d’*. *Behavior Research Methods, Instruments and Computers, 27(1),* 46-51.

Hendricks, J. (1966). Flicker thresholds as determined by a modified conditioned suppression procedure. *Journal of the Experimental Analysis of Behavior, 9*(5), 501-506.

Henson, D.B. (1993). *Visual fields*. Oxford: Oxford University Press.

Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science, 146*, 549-551.

Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 2*(4), 285-302.

Herrnstein, R.J., & de Villiers, P. A. (1980). Fish as a natural category for people and pigeons. In *The psychology of learning and motivation* (ed. G. H. Bower), Vol, 14, pp. 60-97. New York: Academic Press.

Hollander, M., & Wolfe, D.A. (1973). *Nonparametric statistical methods.* John Wiley & Sons, USA.

Honig, W.K., & Urcuioli, P.J. (1981). The legacy of Guttman and Kalish (1956): 25 years or research on stimulus generalization. *Journal of the Experimental Analysis of Behavior, 36(3),* 405-445.

Huber-Eicher, B. (2004). The effect of early colour preference and of a colour exposing procedure on the choice of nest colours in laying hens. *Applied Animal Behaviour Science, 86,* 63-76.

Ikebuchi, M., & Okanoya, K. (1999). Male zebra finches and Bengalese finches emit directed songs to the video images of conspecific females projected onto a TFT display. *Zoological Science, 16,* 63-70.

Iversen, I.H. (1993). Acquisition of matching-to-sample performance in rats using visual stimuli on nose keys. *Journal of the Experimental Analysis of Behavior, 59(3),* 471-482.

Jarvis, J. R., Taylor, N. R., Prescott, N. B., Meeks, I., & Wathes, C. M. (2002). Measuring and modeling the photopic flicker sensitivity of the chicken (Gallus g. domesticus). *Vision Research, 42,* 99-106.

Jitsumori, M., Sieman, M., Lehr, M., & Delius, J.D. (2002). A new approach to the formation of equivalence classes in pigeons. *Journal of the Experimental Analysis of Behavior, 78(3),* 397-408.

Jones, R.B., & Carmichael, N.L. (1998). Pecking at string by individually caged, adult laying hens: colour preferences and their stability. *Applied Animal Behaviour Science, 60,* 11-23.

Jones, R.B., Carmichael, N.L., & Rayner, E. (2000). Pecking preferences and pre-dispositions in domestic chicks: Implications for the development of environmental enrichment devices. *Applied Animal Behaviour Science, 69(4),* 291-312.

Keeling, L.J., & Hurnik, J.F. (1993). Chickens show socially facilitated feeding behaviour in response to a video image of a conspecific. *Applied Animal Behaviour Science, 36(2-3),* 223-231.

Keeling, L. J., & Hurnik, J. F. (1996). Social facilitation acts more on the appetitive than the consummatory phase of feeding behaviour in domestic fowl. *Animal Behaviour, 52(1),* 11-15.

Kendrick, K. M., Atkins, K., Hinton, M. R., Heavens. P., & Keverne, B. (1996). Are faces special for sheep? Evidence from facial and object discrimination learning tests showing effects of inversion and social familiarity. *Behavioural Processes, 38*, 19-35.

Landis, C., & Hamwi, V. (1954). The effect of certain physiological determinants on the flicker-fusion threshold. *Journal of Applied Physiology, 6*(10), 566-572.

Laughlin, S. B. (1981). Neural principles in the peripheral visual systems of invertebrates. In *Handbook of Sensory Physiology,* Vol VII/6B (ed. H. Autrum). Pp. 135-280. Berlin: Springer.

Lea, S. E. G., & Dittrich, W. H. (1999). What do birds see in moving video images? *Current Psychology of Cognition, 18*(5-6), 765-803.

Lewis, P. D., & Morris, T.R. (1998). Responses of domestic poultry to various light sources. *World's Poultry Science Journal, 54*, 7-25.

Lombardi, C.M., & Delius, J.D. (1990). Size invariance in visual pattern recognition by pigeons. In M.L. Commons, R.J. Herrnstein, S.M. Kosslyn & D.B. Mumford (Eds.), *Behavioral approaches to pattern recognition and concept formation.* Hilldale, NJ: Erlbaum, pp 41-65.

Looney, T.A., & Cohen, P.S. (1974). Pictorial target control of schedule-induced attack in white carneaux pigeons. *Journal of the Experimental Analysis of Behavior, 21(3),* 571-584.

Lumsden, E.A. (1977). Generalization of an operant response to photographs and drawings/silhouettes of a three-dimensional object at various orientations. *Bulletin of the Psychonomic Society, 10,* 405-407.

Lundberg, A.S., & Keeling, L.J. (2003). Social effects on dustbathing behaviour in hens: using video images to investigate effect of rank. *Applied Animal Behaviour Science, 81(1),* 43-57.

Madigan, R., & Williams, D. (1987). Maximum-likelihood psychometric procedures in two-alternative forced-choice: Evaluation and recommendations. *Perception & Psychophysics, 42(3),* 240-249.

McKee, S.P., Klein, S.A., & Teller, D.Y. (1985). Statistical properties of forced-choice psychometric functions: Implications of probit analysis. *Perception & Psychophysics, 37(4),* 286-298.

McLean, A.P., & White, K.G. (1983). Temporal constraint on choice: sensitivity and bias in multiple schedules. *Journal of the Experimental Analysis of Behavior, 39,* 405-426.

McQuoid, L. M., & Galef, B.G. (1993). Social stimuli influencing feeding behaviour of Burmese fowl: a video analysis. *Animal Behaviour, 46*, 13-22.

McQuoid, L. M., & Galef, B. G. (1994). Effects of access to food during training on social learning by Burmese red junglefowl. *Animal Behaviour, 48,* 737-739.

Meyer, D.B. (1977). The avian eye and its adaptations. In F. Crescitelli (Ed). *Handbook of Sensory Physiology, Vol VII/5, Visual System in Vertebrates,* Berlin: Springer.

Meyer, D.B., & May, H.C., Jr. (1973). The topographical distribution of rods and cones in the adult chicken retina. *Experimental Eye Research, 17,* 347-355.

Miller, R. J. (1973). Cross-cultural research in the perception of pictorial materials. *Psychological Bulletin, 80*(2), 135-150.

Miloševič, S. (1993). Visual monitoring performance in simultaneous and successive discrimination tasks. *Studia Psychologica, 35(2),* 159-165.

Muñoz-Cuevas, A. (1984). Photoreceptor structures and vision in arachnids and myriapods. In: *Photoreception and vision in invertebrates* (Ed. M. A. Ali). Pp. 335-399, Plenum Press: New York.

Nakagawa, S., Etheredge, R.J.M., Foster, T.M., Sumpter, C.E., & Temple, W. (2004). The effects of changes in consequences on hens' performance in delayed-matching-to-sample tasks. *Behavioural Processes, 67,* 441-451.

Nuboer, J. F. W., Coemans, M. A. J. M., & Vos, J. J. (1992). Artificial lighting in poultry houses: do hens perceive the modulation of fluorescent lamps as flicker? *British Poultry Science, 33,* 123-133.

Ophir, A. G., & Galef, Jr, B.G. (2003). Female Japanese quail affiliate with live males that they have seen mate on video. *Animal Behaviour, 66*, 369-375.

Ord, T. J., Peters, R.A., Evans, C.S., & Taylor, A.J. (2002). Digital video playback and visual communication in lizards. *Animal Behaviour, 63*, 879-890.

Parron, C., Call, J., & Fagot, J. (2008). Behavioural responses to photographs by pictorially naive baboons (Papio anubis), gorillas (Gorilla gorilla) and chimpanzees (Pan troglodytes). *Behavioural Processes, 78*, 351-357.

Patterson-Kane, E., Nicol, D. J., Foster, T. M., & Temple, W. (1997). Limited perception of video images by domestic hens. *Animal Behaviour, 53*, 951-963.

Peissig, J.J., Kirkpatrick, K., Young, M.E., Wasserman, E.A., & Biederman, I. (2006). Effects of varying stimulus size on object recognition in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 32(4),* 419-430.

Pepperberg, I.M., & Wilkes, S.R. (2004). Lack of referential vocal learning from LCD video by grey parrots (Psittacus erithacus). *Interaction Studies, 5(1),* 75-97.

Pepperberg, I.M., Naughton, J.R., & Banta, P.A. (1998). Allospecific vocal learning of grey parrots (Psittacus erithacus): A failure of videotaped instruction under certain conditions. *Behavioural Processes, 42,* 139-158.

Poling, A., Temple, W., & Foster, T.M. (1996). The differential outcomes effect: a demonstration in domestic chickens responding under a titrating-delayed-matching-to-sample procedure. *Behavioural Processes, 36(2),* 109-115.

Powell, R. W. (1967). The pulse-to-cycle fraction as a determinant of critical flicker fusion in the pigeon. *The Psychological Record, 17*, 151-160.

Premack, D. (1976). *Intelligence in Ape and Man*. New Jersey: Lawrence Erlbaum Associates, Inc.

Prescott, N.B., Jarvis, J.R., & Wathes, C.M. (2004). Vision in the laying hen. In: Perry G. (ed) Welfare of the laying hen. CABI Publishing, UK, pp 155-164.

Rieucau, G., & Giraldeau, L. (2009). Group size effect caused by food competition in nutmeg mannikins (Lonchura punctulata). *Behavioral Ecology, 20,* 421-425.

Roget, P. M. (1825). Explanation of an Optical Deception in the Appearance of the Spokes of a Wheel Seen through Vertical Apertures. *Philosophical Transactions of the Royal Society of London, 115*, 131-140.

Rowland, W. J., Bolyard, K.J., Jenkins, J.J., & Fowler, J. (1995). Video playback experiments on stickleback mate choice: female motivation and attentiveness to male colour cues. *Animal Behaviour, 49*, 1559-1567.

Ryan, C.M.E., & Lea, S.E.G. (1994). Images of conspecifics as categories to be discriminated by pigeons and chickens: Slides, video tapes, stuffed birds and live birds. *Behavioural Processes, 33(1-2),* 155-175.

Sacks, R.A., Kamil, A.C., & Mack, R. (1972). The effects of fixed-ratio sample requirements on matching to sample in the pigeon. *Psychonomic Science, 26(5),* 291-293.

Sargisson, R.J., & White, G. (2001). Generalization of delayed matching to sample following training at different delays. *Journal of the Experimental Analysis of Behavior, 75(1),* 1-14.

Schaeffel, F., Howland, H.C., & Farkas, L. (1986). Natural accommodation in the growing chicken. *Vision Research, 26,* 1977-1993.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37(1),* 5-22.

Signal, T.D., Temple, W., & Foster, T.M. (2001). Visual discrimination in the brushtail possum (Trichosurus vulpecula). *Australian Journal of Psychology, 53(3),* 152-154.

Siegel, R. K., & Honig, W. K. (1970). Pigeon concept formation: successive and simultaneous acquisition. *Journal of the Experimental Analysis of Behavior, 13*(3), 385-390.

Shelton, B.R., Picardi, M.C., & Green, D.M. (1982). Comparison of three adaptive psychophysical procedures. *Journal of the Acoustical Society of America, 71(6),* 1527-1533.

Shimizu, T. (1998). Conspecific recognition in pigeons (Columba livia) using dynamic video images. *Behaviour, 135*, 43-53.

Spetch, M.L., & Friedman, A. (2006). Pigeons see correspondence between objects and their pictures. *Psychological Science, 17(11),* 966-972.

Spetch, M.L., Kelly, D.M., & Reid, S. (1999). Recognition of objects and spatial relations in pictures across changes in viewpoint. *Cahiers de Psychologie/Current Psychology of Cognition, 18(5-6),* 729-764.

Spetch, M. L., & Wilkie, D. M. (1994). Pigeons' use of landmarks presented in digitized images. *Learning and motivation, 25(3),* 245-275.

Swaddle, J.P., McBride, L., & Malhotra, S. (2006). Female zebra finches prefer unfamiliar males but not when watching noninteractive video. *Animal Behaviour, 72,* 161-167.

Temple, W., Foster, T.M., & O'Donnell, C.S. (1984). Behavioural estimates of auditory thresholds in hens. *British Poultry Science, 25,* 487-493.

Toda, K., & Watanabe, S. (2008). Discrimination of moving video images of self by pigeons (Columba livia). *Animal Cognition, 11(4),* 699-705.

Trillmich, F. (1976). Learning experiments on individual recognition in budgerigars (Me lopsittacusundulatus). *Zeitschrift für Tierpsychologie, 41(4),* 372-395.

Troje, N. F., Huber, L., Loidolt, N., Aust, U., & Fieder, M. (1999). Categorical learning in pigeons: the role of texture and shape in complex static stimuli. *Vision Research, 39,* 353-366.

Truppa, V., Spinozzi, G., Stegagno, T., & Fagot, J. (2009). Picture processing in tufted capuchin monkeys (Cebus apella). *Behavioural Processes, 82(2),* 140-152.

Ulrich, R., & Miller, J. (2004). Threshold estimation in two-alternative forced-choice (2AFC) tasks: the Spearman Kärber method. *Perception and Psychophysics, 66,* 517-533.

Vaughan, W., & Greene, S. L. (1984). Pigeon visual memory capacity. *Journal of Experimental Psychology: Animal Behavior Processes, 10(2)*, 256-271.

Wasserman, E.A., Kirkpatrick-Steger, K., Van Hamme, L.J., & Biederman, I. (1993). Pigeons are sensitive to the spatial organization of complex visual stimuli. *American Psychological Society, 4(5),* 336-342.

Watanabe, S. (1993). Object-picture equivalence in the pigeon: an analysis with natural concept and pseudoconcept discriminations. *Behavioural Processes, 30,* 225-232.

Watanabe, S. (1997). Visual discrimination of real objects and pictures in pigeons. *Animal Learning and Behavior, 25(2),* 185-192.

Watanabe, S. (1999). How do pigeons see pictures? Recognition of the real world from its 2-D representation. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 18(5-6),* 691-711.

Watanabe, S., Lea, S.E.G., & Dittrich, W.H. (1993). What can we learn from experiments on pigeon concept discrimination? In H.-J. Bischof and H.P. Zeigler (Editors), Avian Vision and Cognition, Cambridge MA, MIT Press, 351-376.

Watanabe, S., Sakamoto, J., & Wakita, M. (1995). Pigeons' discrimination of paintings by Monet and Picasso. *Journal of the Experimental Analysis of Behavior, 63*(2), 165-174.

Watanabe, S., & Troje, N.F. (2006). Towards a "virtual pigeon": a new technique for investigating avian social perception. *Animal Cognition, 9*, 271-279.

Weavers, R. (2000). *Aspects of hens' visual behaviour.* Unpublished Doctorate, University of Waikato, Hamilton.

Weavers, R., Foster, T.M., & Temple, W. (1998). Reinforcer efficacy in a delayed matching-to-sample task. *Journal of the Experimental Analysis of Behavior, 69(1),* 77-85.

Weisman, R. G., & Spetch, M. L. (2010). Determining when birds perceive correspondence between pictures and objects: a critique. *Comparative Cognition and Behavior Reviews, 5,* 117-131.

Werner, C.W., & Rehkämper, G. (1999). Discrimination of multidimensional geometrical figures by chickens: categorization and pattern-learning. *Animal Cognition, 2,* 27-40.

Werner, C.W., Tiemann, I., Cnotka, J., & Rehkämper, G. (2005). Do chickens (Callus gallus f. domestica) decompose visual figures? *Animal Cognition, 8(2),* 129-140.

White, K.G. (1985). Characteristics of forgetting functions in delayed matching to sample. *Journal of the Experimental Analysis of Behavior, 44,* 15-34.

Wyckoff, L.B. (1952). The role of observing responses in discrimination learning, part 1. *Psychological Review, 59(6),* 431-442.

Young, M.E., Peissig, J.J., Wasserman, E.A., & Biederman, I. (2001). Discrimination of geons by pigeons: The effects of variations in surface depiction. *Animal Learning and Behavior, 29(2),* 97-106.

Zeigler, H.P., & Wyckoff, L.B. (1961). Observing responses and discrimination learning. *The Quarterly Journal of Experimental Psychology, 13(3),* 129-140.

Zentall, T.R., Friedrich, A.M., & Clement, T.S. (2006). Required pecking alters judgements of the passage of time by pigeons. *Psychonomic Bulletin & Review, 13(6),* 1038-1042.

Zonderland, J. J., Cornelissen, L., Wolthius-Fillerup, M., & Spoolder, H.A.M. (2008). Visual acuity of pigs at different light intensities. *Applied Animal Behaviour Science, 111*, 28-37.

Zupan, M., Kruschwitz, A., & Huber-Eicher, B. (2007). The influence of light intensity during early exposure to colours on the choice of nest colours by laying hens. *Applied Animal Behaviour Science, 105,* 154-164.