UNIVERSITY OF PADOVA

Department of General Psychology

Bachelor's Degree Course in Psychological Science

Final dissertation

From Visual Discrimination to Violation of Expectation: Developing a visual discrimination task in domestic chicks (*Gallus gallus*)

***Supervisor*: Professor Lucia Regolin**

***Co-supervisor*: Maria Loconsole**

***Candidate*: Svetlana Konovalenko**

***Student ID number*: 2105815**

Academic Year 2025/2026

Contents

Abstract

Violation of expectation (VoE) research so far has primarily focused on corvids and mammals, but other animal models should be investigated as well. The research conducted with domestic chicks (*Gallus gallus*) has scarcely employed the violation of expectation paradigm (VoE), and not ever in combination with a perceptual discrimination task. Hence, for conducting an experiment like this it was necessary to develop and validate a reliable discrimination task procedure. Previous literature provides some evidence of young chickens' perceptual preferences (circular shapes and orange colours) and discrimination abilities (shape, colour and orientation). However, these findings do not necessarily generalise to a novel experimental set-up that would be designed to be used specifically within a VoE paradigm. During the development of the experimental procedure we had to account for the subjects' inherent perceptual preferences, spatial bias, difficulties discriminating stimuli, and occasional loss of motivation. This work illustrates the importance of preparatory studies by reviewing the methodological refinement process that resulted in a functional experimental set-up. We describe and discuss various experimental apparatuses, stimuli sets, reward systems and procedures that were tested, with their strengths and limitations. The procedure presented here has subsequently been tested on a larger sample, validated and adopted for future violation of expectation experiments.

Introduction

1. Developing an experiment

Designing an experimental procedure is a crucial step in scientific research. This process can be complicated by the need to modify and refine the initially planned procedure. A carefully designed experiment not only ensures a valid operationalisation of the research question, but, by saving time and resources for both the experimenters and the subjects, ensures the implementation of key principles of animal research: reduction and refinement. In this work we will discuss the process of creating and validating a visual discrimination task for a future violation of expectation study in baby domestic chicken (*Gallus gallus*).

2. Theoretical background

Violation of expectation (VoE) is a research paradigm originally developed to study pre-verbal cognition in human infants. This paradigm lies in the assumption that looking time, and looking behaviour in general, can indicate interest or surprise. Thus, by measuring the looking behaviour associated with different stimuli we can infer perceptive and cognitive mechanisms of the subject. For instance, in the case of object permanence we observe that infants look longer at a place where they expect a hidden object to reappear.

The VoE paradigm can be especially useful in comparative and animal psychology (Winters et al., 2015). However, violation of expectation research has thus far primarily focused on mammals and corvids, which have been the preferred models since they are generally associated with higher cognitive abilities. The studies on animals also mostly focused on adult individuals, while studies on humans within the VoE paradigm mostly tested children and infants. In this recent overview on the VoE paradigm (Margoni et al., 2023) no studies on chicken are mentioned. Nonetheless, exploring violation of expectation in a precocial bird species could provide novel insights on animal cognition from both an ontogenetic and phylogenetic perspective.

The literature on violation of expectation in chicks is scarce, however, it provides some evidence of possible surprise-like responses in baby chicks. With one study demonstrating the signs of object permanence in young domestic chicks (Szabó et al., 2022) and another showing how an unfamiliar human voice prompted vigilant behaviours in 25-days-old chicks (Bernard et al., 2024). In a study on hens investigating how frustration related behaviour in a discrimination task affects learning (Kuhne et al., 2013) it was demonstrated that after having learned to peck a red or a blue circle (acquisition stage), hens showed significantly more frustration when the correct response suddenly became the opposite of what they just learned (reversal stage). Some of the behaviours associated with the reversal stage can indicate surprise as well as frustration, for example vocalisation, pacing and pecking at the surroundings.

This sets the basis for a future VoE study in baby chicks. Such a study could employ a similar approach to Kuhne's by creating an ethogram of relevant behaviours in violated and non-violated settings to assess surprise or violation of expectation. Among hypothesised surprise-like behaviours would be: looking around, walking around, repetitive checking of the stimuli, distress calls. However, for the data to be reliable it is crucial to design a robust and valid behavioural task. The aim of this work is describing the methodological steps taken in order to develop a visual discrimination task that in the future could be used to study VoE in baby chicks.

3. Visual discrimination in chicks

So far no research on VoE in chicks has been conducted using a visual discrimination task. A first important step in developing such a task is to explore the chick's ability to discriminate between two visual stimuli. Studies of visual discrimination abilities of chicken have provided evidence suggesting that chicks can discriminate between colours, shapes, sizes and, in some cases, orientation.

3.1 Colour discrimination

In the study already mentioned above (Kuhne et al., 2013) the discrimination task was based solely on colour and hens were quick to learn the distinction, needing an average of 61 trials to reach a learning criterion of 18 correct responses in 20 consecutive trials. With younger chicks the training to distinguish colours can become challenging due to innate colour preferences. A study (Salzen et al., 1971) showed that brighter colours like red are preferred by newly hatched chicks. The effect however can be reversed by imprinting, i.e. rearing the subject in social isolation with a blue-coloured disk. In general chicks prefer red to blue, green or yellow (Salzen et al., 1971) and even more so orange over blue (Ham & Osorio, 2007). Choosing a colour that is salient for the baby chick is a good strategy to enhance attention to the presented stimuli, which is why in this project orange was used for most of the stimuli, especially when the task was not colour-based.

No studies investigating if chicks could differentiate between different hues of the same colour were found, that was another dimension for which a set of stimuli was created.

3.2 Shape discrimination

Shape discrimination findings are less robust in chicks compared to colour or size discrimination. It has been shown that chicks struggle with pure shape discrimination tasks (Bingham, 1913), and do not discriminate between a circle and a triangle when size and brightness were controlled for. In a different study young quail chicks successfully discriminated a triangle or a T-shape from two round figures, but not one from another: a disk and a ball - a 2D and a 3D object of the same size (Ono et al., 2002). Based on those findings a shape-based discrimination task could still represent a good solution, as it constitutes a simple visual task that could be solved by chicks, but it requires some methodological attention. For instance, providing additional cues to facilitate discrimination, such as a difference in size or area, to make the task easier overall.

3.3 Orientation discrimination

Another possibility to be considered is whether chicks discriminate stimuli based on their orientation. This was previously investigated in a study: after a short training chicks could discriminate between textures that differ in line orientation (Jones & Osorio, 2004). This finding suggests that an orientation-based task could be another possible strategy for creating a visual discrimination task.

4. Practical considerations and expectations for the final procedure

To conduct a VoE experiment in the future it was not enough to ensure that the discrimination task would be valid, taking into account other factors such as feasibility and ethical considerations was also important. We aimed at creating a task that would not exhaust the chicks with excessively difficult and prolonged training sessions.

We will present these practical considerations and expectations from the discrimination task procedure in a table below (Table 1) together with the rationale behind each one.

Since so many aspects must be accounted for, refining the procedure was important. We will describe how this was done in the following sections focusing in turn on the procedure, the arena, the rewards and the stimuli, discussing the challenges that were faced and the methodological solutions found.

Time constraints	To ensure the subjects' well-being and facilitate future data collection 3-4 subjects should be able to learn the task in one day of training, reaching the criterion in one or two training sessions each no longer than one hour.
Procedure	To ensure training consistency the procedure must be clearly defined, including the number of trials and the order of presentation.
Rewards	To ensure subjects' motivation the optimal rewards must be chosen.
Behavioural response	To observe and record the learning process an appropriate behavioural response should be chosen from the natural behaviour repertoire of young chicks. The training should promote the consistent expression of this behaviour, allowing it to be objectively measured and used as an indicator of the task solved correctly.
Arena	To allow for later VoE testing the set-up should: ● present the stimuli where they can be thoroughly inspected by the subjects; ● hide the reward initially, yet make it accessible when the correct stimulus is chosen; ● allow quick and easy set-up adjustments between trials.

Stimuli	So that in the future VoE study it would be possible to have more than one control group and draw inferences based on the strength (not solely the presence) of expectation – the discrimination task stimuli should have at least two different levels of difficulty.

Table. 1: *Expectations and practical considerations for the discrimination task procedure*

Materials and Methods

1. Subjects and rearing conditions

The subjects in the study were four or five-day-old domestic chicken (*Gallus gallus) and* we tested a total of 33 subjects, out of which 24 were females.

All the experimental activity reported in this work was carried out respecting the applicable national and European laws concerning the treatment of animals in research. Every two weeks the lab would receive fertilized eggs from a local hatchery and incubate them in the laboratory at a temperature of 37.5℃ and humidity of 50%. After hatching, chicks were sexed and housed in groups of three individuals of the same sex in standard metal cages with dimensions of 28 × 32 × 40 cm. Above the cage fluorescent lights of 36 W were placed at a distance of 15 cm. From 7am to 7pm they were on, but from 7pm to 7am the light and the absence of it would be alternating in 2 or 3 hour blocks. At all times the chicks had access to water and food in the cages. Once done with all the experimental activity, usually on day 5 after hatching, the chicks would be donated to local farmers.

2. Procedure

2.1 Testing conditions

For one hour before testing or training the chicks would be food-deprived, to ensure that they would be motivated by food reward, but they would have access to water in their cages at all times. The temperature in the experiment room was kept between 27 and 30 ℃. (On especially hot days, when the temperature outside was 33 ℃ and higher the chick would also have water in the testing arena and could drink at any time to prevent dehydration.) To avoid fatigue and distress in the subjects we kept each training or testing session under 45 – 60 minutes depending on how quickly the subjects got tired, and then would bring them to their cages for a break of at least 30 minutes before another session.

2.2 Waiting

For any training or testing session we would put a chick in a plastic container with a lid and bring it to the experiment room and then after the session back to the cage. In the container the chick could move freely, but we kept a pierced lid on and watched it closely so it would not jump out or run away.

For the Arena 1 set-up the chick would stay in the container while the arena was being prepared for testing or the set-up was being changed between trials. The Arena 1 itself also had a waiting area. When the set-up was ready and all the necessary boxes, the stimuli, and the rewards were placed the chick would spend exactly 3 seconds in the waiting area before each trial to have a chance to focus on the task and the stimuli, then the transparent separator would be removed and the chick would perform the task. After the task was performed or if a mistake was made the chick would go into the container immediately, so it would not see the changes being made to the set-up between the trials, otherwise it would be able to see where the reward was being placed.

In Arena 2 set-up the chick spent the whole experiment session in the arena. The waiting area had a non-transparent separator so it could not see the preparations for the new trial. That improved the training procedure significantly and saved a lot of time. The chick would use the pause between trials to drink if needed, to reorient itself and to get ready for the next trial.

2.3 Preparing the set-up

To begin any training or testing session we would fix the stimuli onto the box lids, the screens or the panel and place the food reward accordingly: the more pleasant reward being associated with the positive stimulus. The positive stimulus would be randomly assigned for each subject.

To avoid bias formation and other training problems we used randomised order to present the positive stimulus on the right or on the left side, with equal number of trials for each one of them and no more than two trials in a row for the same side. Between each trial we would record the answer on paper in the protocol, move the stimuli to their new positions, add the food rewards, and close the lids of the boxes or the drawer to hide the reward. In case of a wrong answer the position would be kept the same and the trial would be repeated either until the correct response was acquired or up to a total of 3 mistakes.

2.4 Pre-training

Before we could test or train the subject we would have to make sure that they would be accustomed to the arena and the set-up first. Later in the process we started calling this “pre-training” and moved it to day 4, as it was more effective that way.

For the initial pre-training we used only one stimulus at a time, the positive one and placed it in the center. The objective was for the subject to inspect the arena freely and get accustomed. In most cases the chick would not approach the stimulus at first, so we would place pieces of worm between the waiting area and the stimulus to reinforce any movement in the direction of the stimulus. Once the subject approached the stimulus without hesitation or pauses we would try not placing any worms on the floor and reward only the pecking of the stimulus, by placing a piece of worm on top of it. It helped for some subjects if during pre-training we kept the lid or the drawer open so they could see the reward inside. The main goal of the pre-training was creating the expectation of certain food next to a certain stimulus and training the subject to complete the correct behaviour once they see the set-up: to move directly and immediately in the direction of the stimulus, to peck on it and to obtain the reward from the food compartment. At this stage they were only presented with the positive stimulus. When this behaviour was repeated for 10 trials in a row with no hesitation the chick was considered ready for the training.

2.5 Training

Before the first training session we would first perform a refreshment of the pre-training to fix the correct response to the positive stimulus and to remind the subject of the set-up and task. Once the behaviour was repeated for three trials in a row with no hesitation the training could start. In case of successful pre-training on the previous day this usually took no more than ten trials and could be done in under five minutes.

Each training session would be video-recorded. The exact starting and ending times, the subject number and sex, the stimuli used would be written down together with the order of presentation and the responses.

During the training the chick was presented with two stimuli at the same time and was expected to choose the positive one. Choosing meant approaching and inspecting the stimulus, pecking on it and receiving the reward. If this happened we moved onto the next trial.

However, we would also observe some unwanted behaviour. For example the chick could refuse to move at all, avoid the stimulus, not peck on it, or make a mistake by choosing the wrong stimuli. In the case of the subject avoiding the task and making no choice we would wait one minute, remove the chick and restart the trial. In all other cases we would remove the subject immediately after the unwanted behaviour occurs and restart the trial.

Throughout the process we used different criteria to determine when the training was complete, but settled on 17/20. This means that if out of any 20 consecutive trials at least 17 ended with correct choices the training was complete.

2.6 End of the experimental session

The training session could last until the chick fulfills the training by reaching the criterion or until it refuses to respond three times in a row, but no longer than 45 – 60 min. After each session we would bring the chick back to the cage and let it rest for at least 30 min before participating in an experiment again. We would do the same if the chicken shows any signs of distress: distress calls, confusion or fatigue. This ensured the well-being and safety of the subjects.

3. Arena

3.1 Arena 1

Arena 1 (see Fig. 1) was an open box made out of plastic plates. A sturdy (1cm thick) plastic square (73 x 73 cm) was used as the base, onto which we glued the walls (3mm thick) in a diamond shape. Arena 1 had a triangular section that served as a waiting area where the chicken was placed behind a transparent plastic piece (25 x 20 cm). The chick spends three seconds there so it would have a chance to focus on the task. Then the separator was removed allowing access to the rest of the arena.

The walls were made 30 cm high so the chick would not be able to jump out. The three walls of the trapeze section were 38 cm tall. The two sides of the triangle were each 62 cm long, the back wall 65 cm, with two smaller walls of 15 cm framing it. In the trapeze part the boxes with stimuli were placed. We centred the boxes, placing them 20 cm away from the side walls and keeping the distance between the boxes at 15 cm to minimise side bias, and ensure both were well perceived.

Fig. 1: *Experimental arena – first version*

3.2 The boxes

The boxes were originally made from polyplast, drinking straws and hot glue (see Fig. 2). They were 5 cm tall, with a rectangle base 5 x 6 cm. The rotating straw allowed for the lid to be pushed down to access the food in a small food compartment inside each box. Without pecking the food could not be seen or accessed. The stimulus was placed on the lid (laminated so it could be easily cleaned) either with double sided tape or by tucking it in a gap in the straw. To hold the lid in place we added two little holders on the sides, made from a thin plastic or paper and taped onto the box.

The boxes had many details and imperfections, since they were made by hand. After some attempts at training we realised that the subjects were basing their choices on how the boxes looked and not on the differences in stimuli. We tried randomising both the position of the stimuli and the boxes, but that did not help.

Fig. 2: *Experimental box – first version*

3.3 Panels

The pecking behaviour was not always easy to train for, so we created two panels (see Fig. 3), comparable in size to the experimental box, to present the stimulus in a way that would allow more natural movement, in this case – walking behind the chosen stimulus to obtain a reward. The panels consisted each of a 6-cm-wide base and 6-cm-tall walls in a trapeze-like shape made out of polyplast and glued together to create a small screen that would hide the reward behind a stimulus.

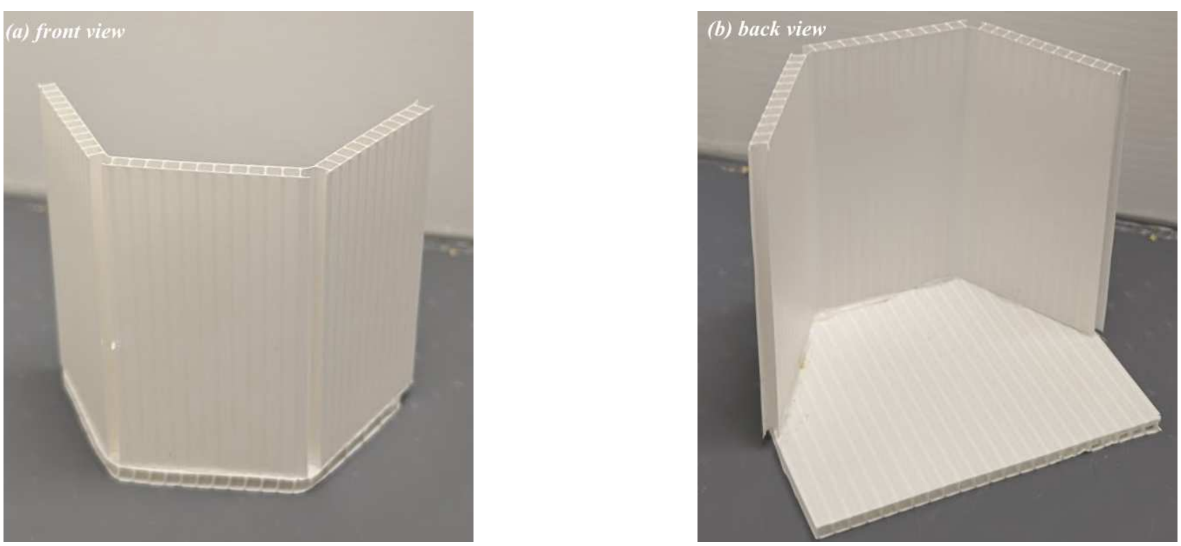

Fig. 3: *Experimental panel*

We used the panels only with one set of stimuli: the orange square set. In comparison to the same task but with the boxes there were no noticeable differences, the training was long and unsuccessful. So we reverted to training with the boxes.

3.4 The 3D printed boxes

The 3D printed boxes (see Fig. 4) had similar dimensions to the ones made earlier from plastic, but they were identical to one another. The only part that could help to tell the boxes apart were the holders added on the sides, so that the lid would stay closed until pushed down. We checked that the chicks would not focus on the holders and would not use them as a distinguishing feature, indeed the holders did not interfere with the experiment.

Fig. 4: *3D-printed experimental box*

The 3D-printed box brought new problems. Since the material was denser, it was hard to find a good way to make and fix the holders – they were either too strong for the chicks to push the lid down or so weak that they became unreliable and the lid would sometimes open without being pushed. Another problem was with the stimuli being held by double sided tape. The material of the box did not hold the tape very well and the stimuli would occasionally detach and fall between or even during trials.

The opening mechanism was similar to the previous one, except now when the lid was pushed down it would hit the food dispenser's back wall making a loud click. We hypothesised that the training was compromised by that and the subjects perceived this click as a reward. To test that theory and fix the problem we created a new arena.

3.5 Arena 2

The new arena (see Fig. 5) was built as a long rectangle (24 x 80 cm), with walls 30 cm tall. A 1-cm-thick hard plastic piece (28 x 100 cm) was used as a base. At 19 cm from the front wall little plastic pieces were glued on so that the separator screen was always set at the same place. The separator-screen was a non-transparent grey plastic piece (22 x 25 cm).

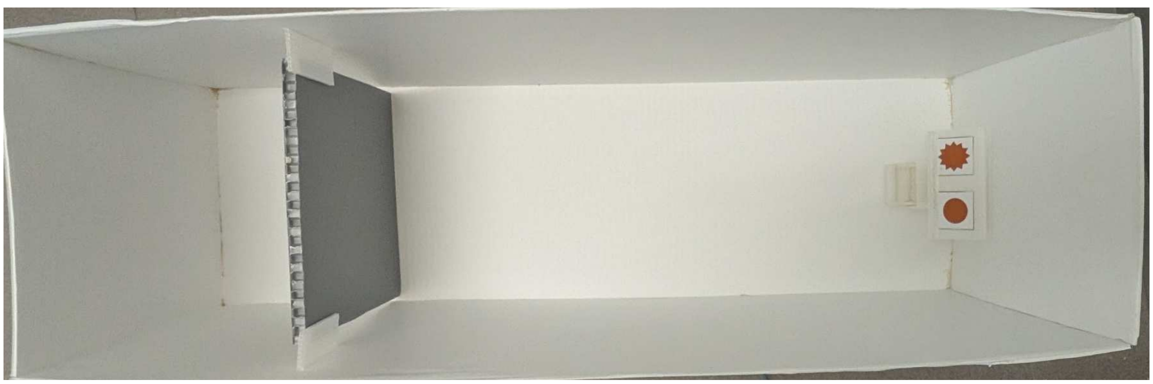

Fig. 5: *Experimental arena – the second version with a panel and a movable food-drawer*

This time the chick could not see the experimental set-up. This was intentional since we would not take the chick out between trials. After a response the chick was gently pushed back to the waiting area, the screen would be placed at the 19 cm mark and the chick had the time and opportunity to reorient itself, while the experimentator replaced the stimuli and maybe added some food reward in the drawer.

At the back end of the arena there was a movable food drawer sized 5,5 x 5 x 3 cm with a panel sized 11,5 x 6cm above it. The drawer was pushed forward or pulled backward by a straw, the subjects however would not see the experimenter's hand controlling the movements of the drawer. This ensured precise manual control over the reward presentation. Both stimuli would be placed on the panel (see Fig. 6). This fixed the problem of the subjects accidentally overlooking one stimulus or another – here they were exhibited at a suitable height and could be viewed together.

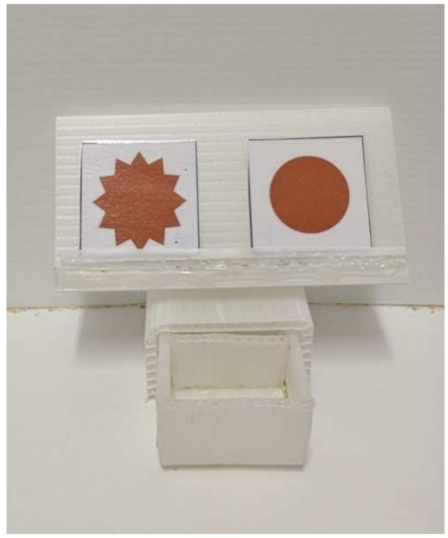

Fig. 6: *Panel and movable food drawer, in the Arena 2 experimental set-up*

4. Rewards

Throughout the experimentation process we have used different rewards to motivate the subjects: mealworms and grain. “Grain” was the main component of the chicks regular diet: a special grain mix that contains all the necessary nutrients for growing chicken.

As a stronger reward associated with the positive stimuli we used their preferred food: mealworms, the larvae of *Tenebrio molitor* (in this work often referred to as ‘worms’). This section describes different reward systems used during training in chronological order.

It is important to note that for the pre-training, especially when just getting accustomed with the set-up, chicks received small pieces of worms, about 2 – 3 mm in length, freshly cut and placed on the floor, on the box lid or on the stimulus to make them explore and peck. Then if the wanted behaviour (such as pecking on the stimulus for example) persisted even without this additional reward the pre-training was considered successful and from that point on the rewards would only be delivered for the correct completion of the task.

4.1 All-or-nothing approach

Our first strategy was all-or-nothing: pecking on the positive stimulus would open the box, from which the chick could retrieve half a worm and eat it; while pecking on the wrong stimulus meant no reward and immediate end of the trial. Some subjects tried to open both boxes during the same trial. We prevented that by removing the subject from the arena.

However there were subjects that were very fast and hard to catch and some managed to see inside the other box, but they showed no improvement in training: the wrong answers persisted even after discovering the absence of the food reward.

We have also tried training with grain as the main reward associated with the positive stimulus, but subjects showed very little motivation and refused to perform the task after just a couple of trials.

The all-or-nothing rewards system was functional but had some problems: chicks got full before the experimental session was over and lost interest in the task, after several incorrect responses, they would not open any box.

To explain why wrong answers persisted even if the chick saw that the box with the wrong stimulus was empty we analysed the set-up. The click of the box lid when it fell down might have been the problem. We hypothesised that the click itself acted as a reward, since already during pre-training he chicks were conditioned to associate that click with the reward inside. This was one of the main reasons why a completely new set-up was developed and we have built the Arena 2 with the movable drawer.

4.2 Unexpected reward

To avoid the subjects getting full too fast in Arena 2 we started the training with grain as the main reward. To make the reward accessible we filled the drawer full of grain. Three pecks of food reward were allowed before removing the chick from the drawer and closing it. We observed low motivation, so we started hiding small (2 – 3mm long) pieces of worm in the grain. This way the chicks associated the worm with the positive stimulus, but did not get it every time, keeping their motivation high. In case of a mistake, no response or other unwanted behaviour the drawer would not open, the chick would be removed and the trial would start over. With the manual control of the food drawer we were able to significantly reduce the number of accidents and mistakes in providing the food reward.

4.3 Additional reward

Considering the insight about the clicking lid of the box from Arena 1 we started using a dog clicker in training. The clicks would be introduced already at pre-training and multiple favorable pecks would be followed by multiple clicks. This was a successful strategy and it significantly increased the productivity of the training sessions, as well as shortening them. The combination of small pieces of worm hidden in the grain together with a dog clicker was adapted as the final reward system for future discrimination training.

5. Stimuli

During the development of the experiment we have tried a lot of different sets of stimuli to find a suitable one. The stimuli had to be such that the chicks could learn to tell them apart, but would not show a strong preference towards one or another. A preference towards one would mean easy or effortless training for subjects which have the preferred stimulus as the positive, but at the same time it would make the training of the other group longer, harder or even impossible, while training on the stimuli that are naturally less appealing.

The stimuli also had to have different levels of difficulty. In the sets created for this experiment we called the two stimuli that were easier to tell apart "level 1 pair" and the other pair "intermediate stimuli". Most of the time we tested only the level 1 pair, since moving onto another difficulty level was only done when the easy level was passed.

To prove that the stimuli pairs created by the experimenters would be actually perceived as more or less difficult by the chicks we would have to observe different training times, with more time spent training for the intermediate stimuli.

To create a set of stimuli for a discrimination task a pair of simple images was used. The images would differ in one certain aspect: size, colour, orientation or shape. This section presents the pairs of stimuli we used and have tested to find the one to satisfy all of those demands.

5.1 Stimuli different in size

We started with a set of orange circles of different sizes placed in the center of 40 x 40 mm white squares. The colour orange was chosen because the newly hatched chicken generally have been shown to like it, with this preference persisting up to when they are 9-days-old (Ham & Osorio, 2007) and the circle was chosen since chicks show preference to circular shapes (Fantz, 1957). This way we wanted to make the stimuli look more appealing, using preferred shape and colour to ensure that the subjects would not be scared and would interact with the stimuli.

The level 1 pair consisted of a 7 mm circle and a 27 mm one (see Fig. 7). The intermediate stimuli for this set were never used.

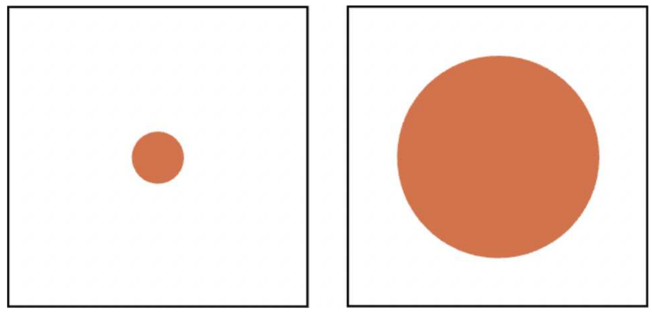

Fig. 7: *Small-big circle – level 1 pair*

This stimuli pair was not successful in training. One possible reason might be that the chicks already had a strong preference towards the smaller shape. It was noted before (Goodwin & Hess, 1969) that smaller shapes elicit a higher level of pecking response. Since our training relied heavily on pecking response, this preference compromised the training.

5.2 Stimuli different in colour

It was decided to simplify the task and try out some stimuli of two different colours. So the next set consisted of a white and a black square also 40 mm wide (Fig. 8). The idea for the intermediate stimuli was to use different shades of grey, however they were not tested.

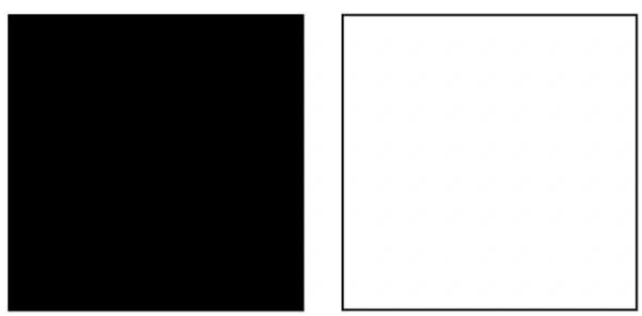

Fig. 8: *Black-white square – level 1 pair.*

The training on the black square as a positive stimulus was effective, the criterion of 10/10 was met by the subject immediately, but training the other subjects on the white square was never completed. The reason for that was that even though the white square had a visible outline it was not perceived by the chicks as a stimulus, it seemed an empty space to them. The subjects behaved as if there was no stimulus at all, refusing to peck on it.

5.3 Stimuli: different hues of a the same colour

The next attempt was done with the set of orange squares of different hues. We used 2 pairs of squares: one was very easy to discriminate (Fig. 9), the other much harder (Fig.10), sometimes even the experimenters had problems telling them apart.

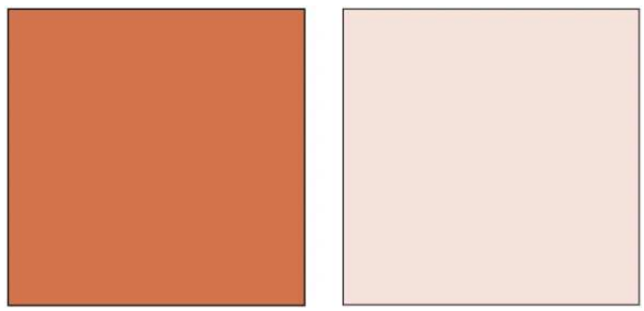

Fig. 9: *Orange square set – level 1 pair*

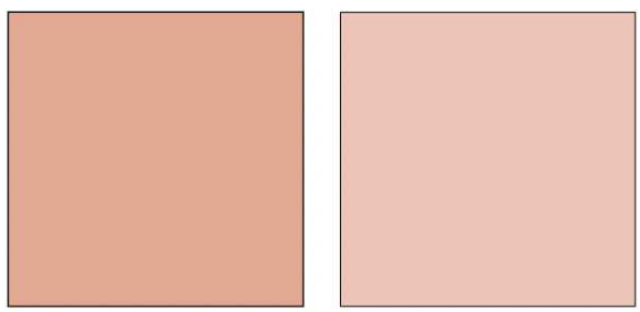

Fig. 10: *Orange square set – intermediate stimuli*

The training with this set was slow, with accuracy level between 54 – 85%, out of nine subjects only one chick passed the criterion. We trained one other chick directly on the intermediate stimuli: the criterion was not reached and the whole stimuli set was discarded as too difficult.

5.4 Stimuli presented with a different orientation

To test for orientation discrimination a set of stimuli was created containing an upper case letter A and an upside down A of the same shape. It was striped with alternating green and purple stripes. The intermediate stimuli would be created by removing the green stripes making the stimuli incomplete.

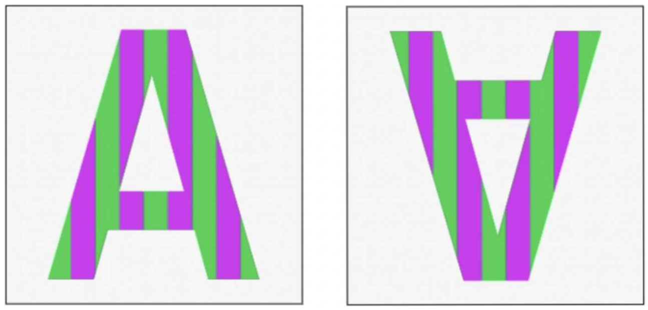

Fig. 11: *Upside down letter set – level 1 pair (A – upside down A)*

In order to use this set we had to move the boxes and place them directly next to the wall, so it would be impossible to go around the box. Otherwise the subjects inspecting the stimulus from different sides would see the letter upside down even when it was not meant to be perceived as such.

The training on these stimuli was difficult. Two out of three subjects tested had developed a strong side bias that was not possible to correct even by changing the position of the arena or the experimenter. The number of correct responses stayed below chance level.

5.5 Stimuli different in shape with similar local features

Trying to improve the letter set we added the letter N and tested the A and N pair next.

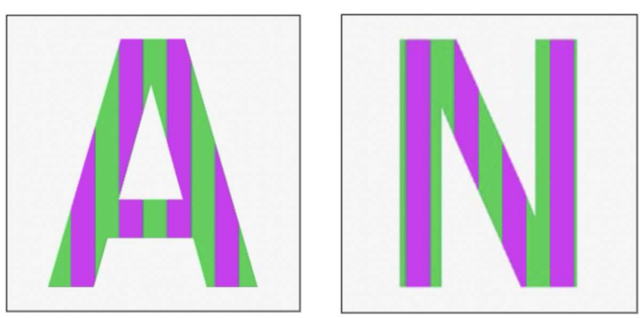

Fig. 12: *Letter set – level 1 pair (A – N)*

The letter N shares some local features with the letter A, like angles, yet has a completely different overall shape. The training on this pair was unsuccessful, the side bias was adopted by the subjects as their main strategy. In the attempt to correct for it we changed the protocol: instead of waiting for 3 wrong answers in a row to change the position of the positive stimulus, we started switching the positions of the stimuli each trial independently of the answer. That was not helpful and this set was discarded as too difficult.

5.6 Stimuli different in shape: the star-circle set

At that point the training continued in Arena 2 with the movable drawer. The small-big circle set was tested once again (see section 5.2 and Fig. 7), but it was not successful: the chicks preferred the small circle to the big one and the accuracy stayed below chance level.
So a different shape-based stimuli set was created: the star-circle set that consisted of a big orange circle (27 mm) and stars with 4, 5, 6, 12 or 32 spikes. The level 1 pair (Fig. 13) consisted of a 4-spiked star and a circle, the intermediate stimuli (Fig. 14) – of a 5-spiked star and a 32-spiked star.

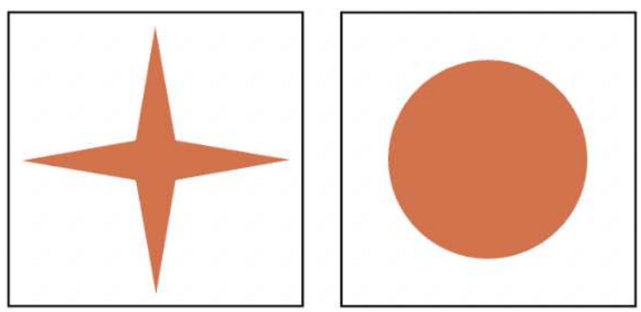

Fig. 13: *Star-circle set – level 1 pair*

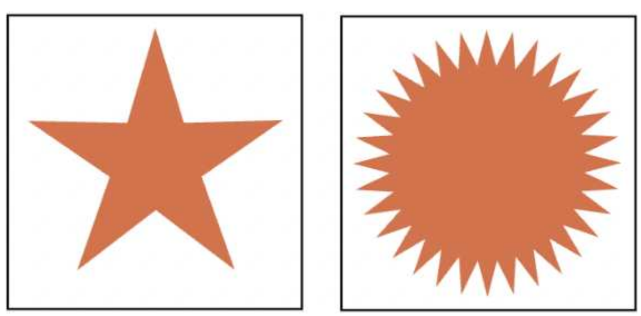

Fig. 14: *Star-circle set – intermediate stimuli*

The training was taking two or three sessions and more than a hundred trials in total before reaching the criteria. Even though some subjects reached the criterion the training time was too long. Another challenge was determining which stimuli should be paired together and whether a 5-spike star or a 32-spike star would be harder to tell apart from a circle.

5.7 Stimuli different in shape: the triangle-circle set

We moved on to the last set containing triangles and circles (each figure being 2 cm in width centered on a 3-cm-wide white square). An orange circle and an orange triangle were used as level 1, and for the intermediate stimuli we used a triangle with the corners rounded up by 4 mm and another one by 6 mm.

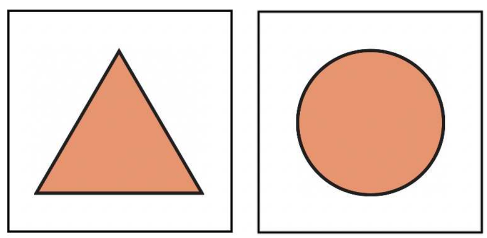

Fig. 15: *Triangle-circle set – level 1 pair*

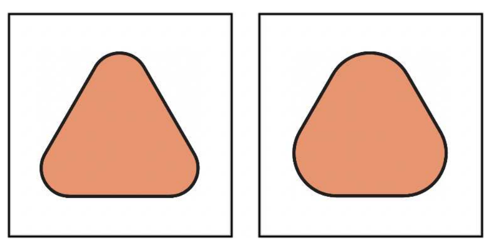

Fig. 16: *Triangle-circle set – intermediate stimuli*

This set was successful in training: all tested subjects had completed the pre-training. Three out of three completed the training, meeting the criterion on the level 1 discrimination task and one chick meeting the criterion on the intermediate stimuli. The fact that not all subjects were able to complete the training for the intermediate stimuli and the one that did needed more trials and time showed that they were indeed perceived as more difficult. The results did not differ for the subjects with different positive stimuli, ensuring there was no strong preference for the circle or the triangle in this case. Since it fulfilled all of the necessary requirements it was this set that was chosen for the discrimination training and testing.

Discussion

The aim of this work was developing a simple task to be employed in the future testing of a complex cognitive ability in baby chicks using the violation of expectation paradigm. We have designed and refined a training procedure with a visual discrimination task and a shorter summary of those changes and methodological choices can be found in a table at the end of this section (Table 2).

After trying different approaches in training, the procedure was well defined, with randomised balanced order of stimuli presentation, no more than two consecutive trials on the same side and the use of unexpected and additional rewards. The training with the final procedure took an appropriate amount of time: one to two short pre-training sessions on day 4 and one to three sessions on day 5 to learn the discrimination of one pair of stimuli.

It was challenging in the beginning to understand which parts of the procedure were not functioning properly, as there were many aspects where the subject's learning difficulties might have stemmed from. This was solved with experimentation and creativity, as well as by making one change to the procedure at a time. Even though some difficulties could not be solved and were based on external factors, e.g. switching the genetic line made some subjects too heavy to want to move on day 5.

Different experimental set-ups were tested, starting with Arena 1, where the chicks were presented with the stimuli fixed onto two boxes and were trained to peck to open the lid and access the food. The reasons why Arena 1 with the boxes did not work as intended were not fully understood. One issue might be lying with the simultaneous discrimination difficulties chicks have demonstrated in a study by (Vallortigara & Zanforlin, 1988), where it was found that coloured disks added on top of the boxes did not aid the learning, since the subjects were more heavily relying on the relative spatial position. The authors noted that simultaneous discrimination, i.e. discrimination based on both the colour and the box position, was difficult for the chicks. This might explain why in the Arena 1 the side biases persisted: probably inspecting and comparing both stimuli on two different boxes was too difficult and chicks reverted to a simpler strategy and relied on spatial location to solve the task.

In the Arena 2 this problem was solved by presenting both stimuli closer together on the same panel. The set-up with the movable drawer allowed for a controlled and efficient way

to deliver the rewards. Another notable improvement was that the time to switch stimuli positions and adjust the set-up was less with Arena 2, due to its structural simplicity.
The reward system as well has undergone some changes. Training on mealworms made the subjects full quickly and lose motivation before making progress in learning the task. Hiding the worm pieces in the grain was found to be the best strategy to ensure motivation by providing the preferred food as the reward, but making it unpredictable and scarce. Adding the clicks provided additional reward for additional pecks, which in turn helped boost learning.
For what concerns the visual stimuli we have tried colour, colour hues, size and orientation-based stimuli, none of which worked as well as the final shape-based task with the triangle-circle set. The main difficulties with the stimuli were creating a task that would be easy enough for our subjects and defining the difficulty levels. However, those issues were resolved and with the triangle-circle set we could demonstrate both quick and efficient learning of the level 1 stimuli (all three subjects passed the criterion) and the difficulty of intermediate stimuli (only one subject passed the criterion and after significantly more trials).

Time constraints	The last three subjects tested were able to learn the task in one day of training, reaching the criterion for level 1 stimuli within one or two training sessions each no longer than one hour
Procedure	The final procedure was clearly defined and standardised, including the number of trials, order of presentation and rewards system. It has already been successfully adopted by other experimenters working on validating the task on a bigger sample
Rewards	An optimal reward system was chosen which facilitated and boosted learning by providing unexpected and additional rewards (hidden pieces of worm in the grain and clicks).

Behavioural response	Pecking was chosen as a behavioural response; by refining the set-up and pre-training it became possible to train for it in less than ten trials. In addition to that the task allowed for recording of secondary behaviours, such as latency to approach the stimuli, vocalisation or pecks on something other than the stimuli.
Arena	In the Arena 2: ● the stimuli were presented together at a suitable height, allowing thorough inspection and comparison; ● the reward was initially hidden in the movable drawer but appeared when the correct stimulus was chosen; ● quick and easy set-up adjustments between trials;
Stimuli	The triangle-circle set chosen for the discrimination task has two different levels of difficulty, with subjects needing more time and trials to learn to discriminate between the intermediate stimuli.

Table. 2: *Discussion of the final discrimination task procedure*

Conclusion

The achieved protocol with the training conducted in Arena 2 on the triangle-circle stimuli set was adapted for further experiments on violation of expectation. The procedure has since been validated, as described in this work, and will soon be used to test chicks for violation of expectation. The challenges and vast variety of different methodological solutions tested illustrate how important preparatory studies are to scientific research. Even though the study did not directly investigate surprise response in baby chicks, it sets the basis for further studies that can investigate this research question exploiting the proposed methodology.

References

1. Barber, T. A. (2016). Discrimination of shape and size sues by day-old chicks in two one-trial learning tasks. *Behavioural Processes*, *124*, 10-14. https://doi.org/10.1016/j.beproc.2015.11.019
2. Bernard, J., der Laan, M.-H. P., Calenge, F., Lecoeur, A., Calandreau, L., Dutour, M., Même, N., Lansade, L., Seressia, J., & Ferreira, V. H. B. (2024). Discrimination of familiar and unfamiliar human voices is independent of prolonged human-animal interaction in domestic chicks. *Applied Animal Behaviour Science*, *270*. https://doi.org/10.1016/j.applanim.2023.106138
3. Bingham, H. C. (1913). Size and form perception in Gallus domesticus. *Journal of Animal Behavior*, *3*(2), 65-113. https://doi.org/10.1037/h0075283
4. Fantz, R. L. (1957). Form preferences in newly hatched chicks. *Journal of Comparative and Physiological Psychology*, *50*(5), 422–430. https://doi.org/10.1037/h0044973
5. Goodwin, E. B., & Hess, E. H. (1969). Innate Visual Form Preferences in the Pecking Behavior of Young Chicks. *Behaviour*, *34*(4), 223-237. https://doi.org/10.1163/156853969X00134
6. Ham, A. D., & Osorio, D. (2007). Colour preferences and colour vision in poultry chicks. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1621), 1941–1948. https://doi.org/10.1098/rspb.2007.0538
7. Jones, C. D., & Osorio, D. (2004). Discrimination of oriented visual textures by poultry chicks. *Vision Research*, *44*(1), 83-89. https://doi.org/10.1016/j.visres.2003.08.014.
8. Kuhne, F., Sauerbrey, A., & Adler, S. (2013). The discrimination-learning task determines the kind of frustration-related behaviours in laying hens (Gallus gallus domesticus). *Applied Animal Behaviour Science*, *148*(3-4), 192-200. https://doi.org/10.1016/j.applanim.2013.09.003
9. Margoni, F., Surian, L., & Baillargeon, R. (2023). The Violation-of-Expectation Paradigm: A Conceptual Overview. *Psychological Review*, *131*(3), 716-748. https://www.researchgate.net/publication/373114262_The_Violation-of-Expectation_Paradigm_A_Conceptual_Overview

10. Ono, Y., Hayashi, I., & Matsushima, T. (2002). Visual memory of shapes in quail chicks: discrimination among 2-dimensional objects. *Zoolog Sci*, *19*(7), 719-25. doi: 10.2108/zsj.19.719. PMID: 12149571
11. Salzen, E. A., Lily, R. E., & McKeown, J. R. (1971). Colour preference and imprinting in domestic chicks. *Animal Behaviour*, *19*(3), 542-547. https://doi.org/10.1016/S0003-3472(71)80109-4.
12. Szabó, E., Chiandetti, C., Téglás, E., Versace, E., Csibra, G., Kovács, Á. M., & Vallortigara, G. (2022). Young domestic chicks spontaneously represent the absence of objects. *eLife*, *11*. https://doi.org/10.7554/eLife.67208
13. Vallortigara, G., & Zanforlin, M. (1988). Simultaneous Discrimination Learning in Chicks: Spatial Representations and Object Characteristics. *Ethology*, *79*, 248-256. https://doi.org/10.1111/j.1439-0310.1988.tb00714.x
14. Winters, S., Dubuc, C., & Higham, J. P. (2015). Perspectives: The Looking Time Experimental Paradigm in Studies of Animal Visual Perception and Cognition. *Ethology*, *121*, 625-640. https://doi.org/10.1111/eth.12378