Analysis of Learning of an Echolocation Task in a Bottlenose Dolphin

A thesis submitted towards the partial fulfilment of BS-MS Dual Degree Programme

by

Vishruta Prashant Yawatkar

(20131102)

For the study conducted under the guidance of

Prof. Mandar Chitre

and

Dr. Matthias Hoffmann-Kuhnt

at the

Acoustic Research Laboratory, Tropical Marine Science Institute,

National University of Singapore, Singapore

Certificate

This is to certify that this dissertation entitled "**Analysis of Learning of an Echolocation Task in a Bottlenose Dolphin**" towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by **Vishruta Prashant Yawatkar** at the Acoustic Research Laboratory, Tropical Marine Science Institute, National University of Singapore, Singapore under the supervision of **Prof. Mandar Chitre, Associate Professor, Department of Electrical and Computer Engineering, National University of Singapore** and **Dr. Matthias Hoffmann-Kuhnt, Senior Research Fellow, Acoustic Research Laboratory, National University of Singapore** during the academic year 2017-18.

1. Prof. Mandar Chitre

2. Dr. Matthias Hoffmann-Kuhnt

Signature of Student

Signature of Supervisor(s)

Declaration

I hereby declare that the matter embodied in the report entitled "**Analysis of Learning of an Echolocation Task in a Bottlenose Dolphin**" are the results of the work carried out by me at the Acoustic Research laboratory, Tropical Marine Science Institute, National University of Singapore, Singapore under the supervision of **Prof. Mandar Chitre** and **Dr. Matthias Hoffmann-Kuhnt** and the same has not been submitted elsewhere for any other degree.

1. Prof. Mandar Chitre

2. Dr. Matthias Hoffmann-Kuhnt

Signature of Student

Signature of Supervisor(s)

Contents

Abstract

Echolocation is one of the important senses in dolphins that contributes to their perception of the underwater environment. However, many questions still remain unanswered with respect to the development of their echolocation sense. It is observed that dolphins can be trained quite fast to perform an object recognition task. However, it is unknown whether dolphins have to learn to echolocate after birth or have the innate ability to do so. Here, an indirect approach of studying the acoustics in a non-neonate dolphin was conducted, to understand the development of echolocation in dolphins. A dolphin, named Angelo, was trained at a young age to perform several cross-modal matching tasks. These tasks involved echolocating an object concealed inside an anechoic box, followed by having to match it visually to the same object among several alternatives presented in air. The analysis of the learning curves associated with the dolphin's progress showed that Angelo understood the concept of matching-to-sample. Along with the progress of the training, Angelo was observed to fine-tune some acoustic parameters in the clicks he transmitted. Both the average click number used by Angelo across trials and the variance of the mean frequencies of clicks in individual trials decreased as the training progressed. This fine-tuning of the acoustic parameters by Angelo suggested that there was learning of echolocation involved.

List of Figures

List of Tables

Acknowledgements

I would like to express my sincere gratitude to Prof. Mandar Chitre and Dr. Matthias Hoffmann-Kuhnt for their guidance. They were very supportive throughout the project.

I am thankful to Dr. Raghav Rajan for his comments during the course of the project.

It would had not been easy to complete the work without the suggestions of Dr. Hari Vishnu and Abel Ho. I would also like to thank Dr. Ahmed Mahmood and Dr. Chong Wei for their helpful comments. I am grateful to Angelo, the dolphin for performing the trials. The project would had not been possible without the efforts of Eszter Matrai, Rick Kwok and all the interns at Ocean Park who conducted the trials. I also thank my family for their constant support and faith in me.

Lastly, I am thankful to the Acoustic Research Lab (Tropical Marine Science Institute, National University of Singapore) and Ocean Park, Hong Kong for the support and funding for the project.

Introduction

One of the primary senses that dolphins use to investigate their environment is echolocation. Since sound travels faster and further in water than on land as compared to light, echolocation is an essential way for cetaceans like dolphins to sense and interpret their surroundings, apart from using their visual sense (Au, 2000). On land, animals like bats, birds (e.g., swiftlets and oilbirds) and shrews use echolocation as well (Griffin, 1950; Pye, 1960; Griffin, 1953; Nowacik, 1959; E. Gould et al. 1964). However, echolocation is comparatively rare in animals, and there are still many aspects of it which are unknown. For example, in dolphins, not much is known about the sensory development of the sense of echolocation.

Studies have been performed to investigate the development of echolocation in neonate bottlenose dolphins which are a few weeks old (Harder et al., 2016). However, in order to understand if the ability to echolocate in neonate bottlenose dolphins is innate, their acoustic signals would have to be studied as soon as they are born. The new born dolphin calves are vulnerable without their mother thus making it difficult to separate them. This makes it complicated to distinguish the acoustic signals specific only to the new born calves. Thus, an indirect approach could be employed in non-neonate dolphins in order to understand their development of echolocation.

Along with the echolocation ability, vision is also well developed in dolphins (Herman et al., 1975). In bottlenose dolphins cross-modal perception using echolocation and vision has been investigated (Herman et al., 1998). The cross-modal matching ability of dolphins can be used in order to investigate shape recognition through echolocation. The term cross-modal matching indicates a perception that involves interaction between two sensory modalities. If the ability of echolocating on an object is not completely innate but has to be learned by the dolphin during the course of the training, some of the acoustic parameters of the echolocation signal might undergo changes as the dolphin learns to fine-tune its ability to echolocate in order to identify the object correctly and pick the matching alternative. As opposed to this, if the ability to echolocate were innate, the dolphin would be expected to only develop its understanding during the training of the task and not fine-tune its physical parameters much.

MTS Training for the bottlenose dolphin, Angelo

With the hypothesis as detailed above, data from the set of experiments that had been conducted from September 2013 to March 2015 at Ocean Park, Hong Kong were analysed. These experiments involved training a young male bottlenose dolphin named Angelo to perform a cross-modal matching or matching-to-sample task (MTS task). They consisted of different types of trials that required Angelo to perform matching, starting with using only the visual perception to identify both the sample object and match it with one of the alternatives (visual-to-visual or VV task). In a visual-to-visual task both the sample object and the alternatives were presented in air. Later on, the trials involved the use of echolocation to identify the sample object from a box covered with opaque Plexiglas that was placed underwater - and the visual perception to identify the alternatives presented in boxes placed in air (echoic-to-visual or EV task). The number of alternatives from which Angelo had to choose the correct object, started with two and then went on to include up to four alternatives during the course of the training. The training began with the visual-to-visual tasks using the basic training objects. Later, PVC objects were introduced both for the visual-to-visual and echoic-to-visual tasks (Table 1) (Figure 4 and Figure 5).

Sets of Objects Used	Object codes	Object Type
Training Objects	YS, FP, AI, KB, WC, SV	Plastic toys
Baseline Objects	FF, SQ, OP, DL	PVC object
Set 1	EL, LR, HH, KY	PVC object
Set 2	OC, BM, CO, RI	PVC object

Table 1: Objects used for the matching-to-sample tasks
Training for Angelo started with the training objects. The baseline objects were then introduced followed by the set 1 and set 2 objects. The echoic-to-visual training began with the Baseline objects.

The setup for the visual-to-visual trials changed as the training for Angelo progressed in the initial phase. In the first month of training, Angelo was learning to touch the training objects and then match them. Initially both the sample object and the alternatives were presented at close locations. By the end of the third month of training, in November 2013, Angelo was performing the visual-to-visual MTS task in the final setup in which the sample object and the alternatives were presented at different locations.

In case of the echoic-to-visual trials, the training for echoic matching started with the habituation to the presence and the use of a single anechoic sample box or E box. This box was later used for the presentation of the sample stimuli. Angelo was reinforced for swimming close, and later for stationing in front of the sample box. He was reinforced to follow a target behind the opaque sample box and monitored for indication of successful target localisation. These sessions were referred to as MSHAB (Habituation) sessions; they continued even after the echoic-to-visual trials had started.

In order to perform a MTS task, Angelo would approach from a distance towards the sample box. For the echoic-to-visual task the sample box was underwater (Figure 1). Angelo would station in front of the sample box in order to echolocate on the sample object then turn away to choose an object from the alternatives by pressing on the paddle (Figure 2). The alternative boxes had blinds covering the stimuli which were opened at the same time for all the boxes. This ensured that Angelo was seeing all the alternatives at once and only after he had perceived the sample object.

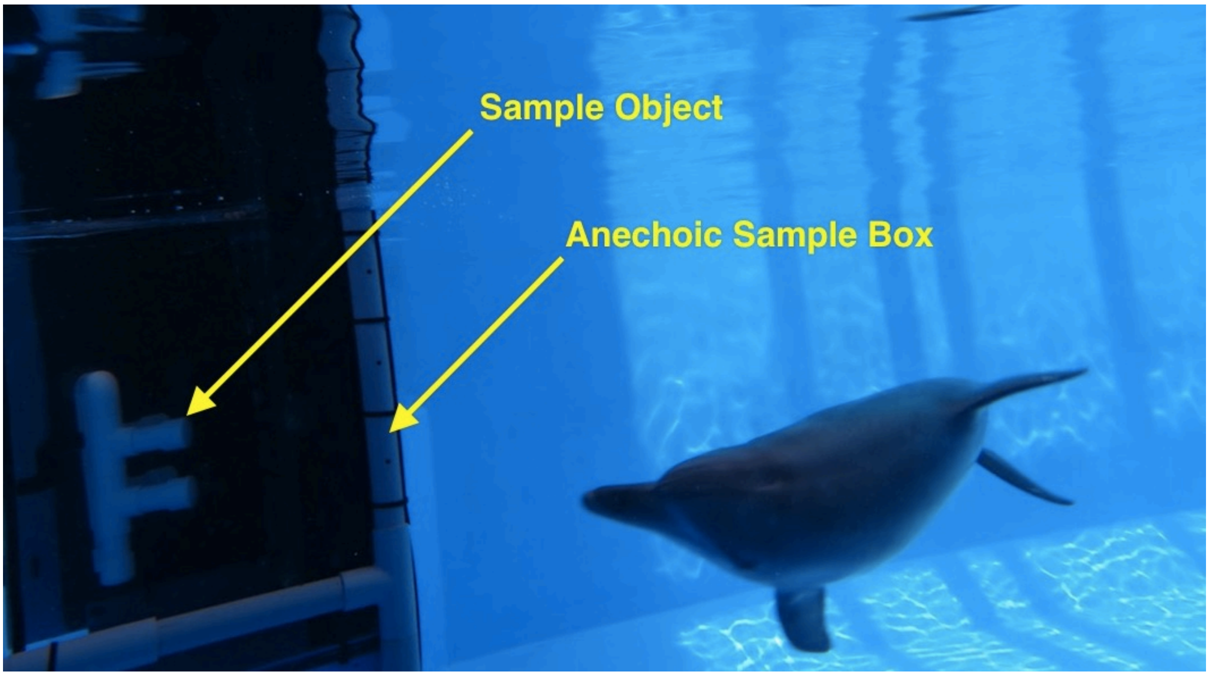


Figure 1: Angelo echolocating on a sample object
In an echoic-to-visual task, Angelo was presented the sample object underwater which was concealed in an anechoic sample box mounted with hydrophones in order to record the acoustics.

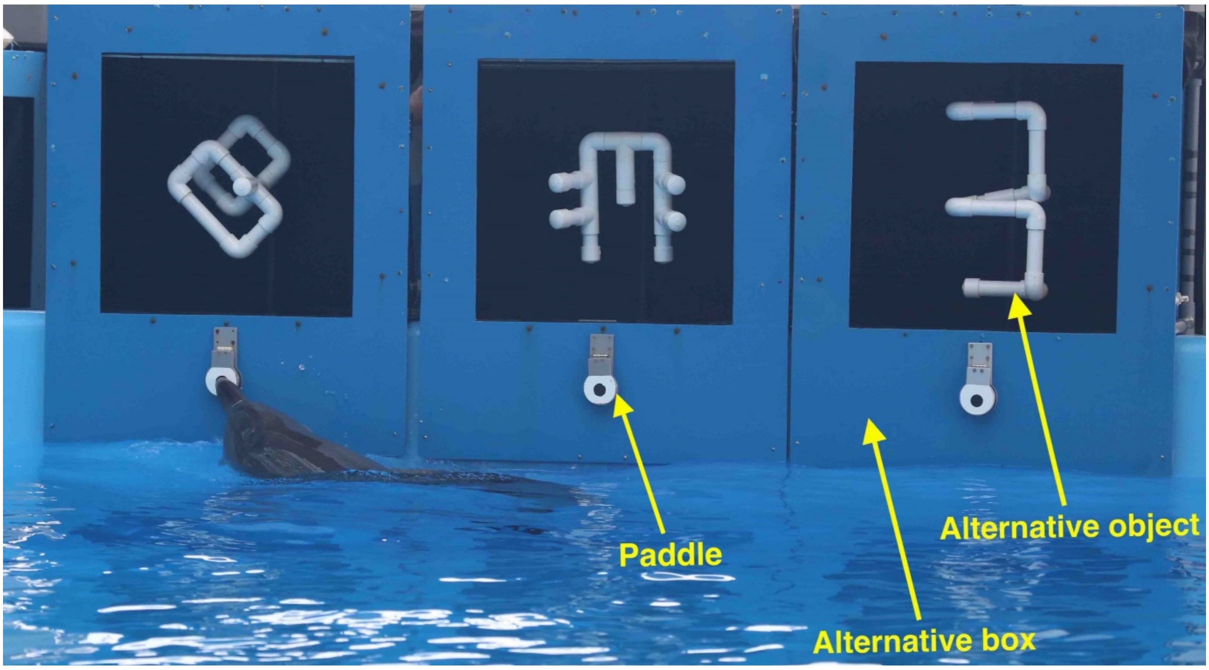


Figure 2: Angelo choosing from three alternatives
In both the visual-to-visual and echoic-to-visual MTS tasks Angelo had to choose from alternatives presented in boxes placed in air by pressing a paddle. The number of alternatives for a trial ranged from two to four.

Training with partially covered Plexiglas

The training with the partially covered Plexiglas was performed during October 22, 2014, to November 13, 2014. During this training the sample box was not completely sealed. It was completely uncovered at the beginning, and eventually, the coverage was increased with increments of 15 cm until it was again completely covered on November 13, 2014 (Figure. 3). This training was performed in an attempt to improve Angelo's understanding of the echoic-to-visual tasks.

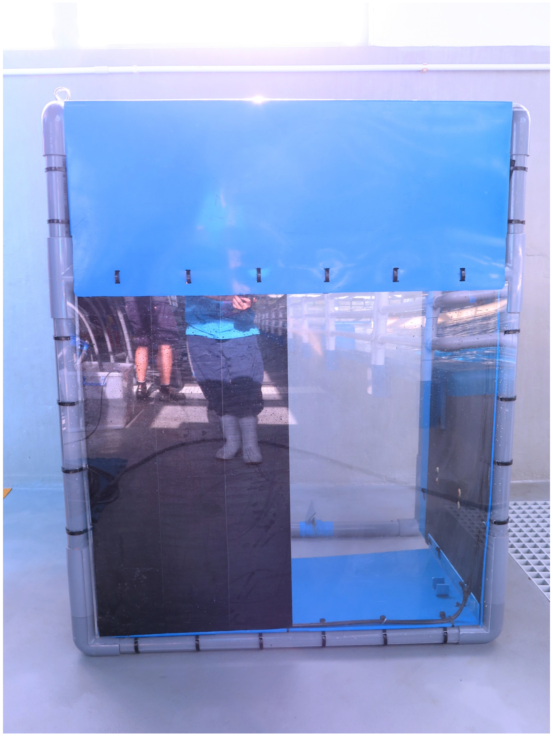


Figure 3: Sample box with partially covered Plexiglas
The sample box (placed outside water) is covered partially with the Plexiglas (in black), the rest is clear. This view is similar to Angelo's view of the partially covered sample box underwater. The coverage of the Plexiglas was increased in the increments of 15 cm.

A pair of baseline objects, FF and SQ, was used in the initial echoic-to-visual trials. The trials with the FF and SQ used as sample objects were highest among trials for all the other objects followed by another pair of baseline objects OP and DL (Figure. 4). The Set 1 and Set 2 objects were introduced after a considerate number of trials had been performed using the baseline objects (Figure. 5).

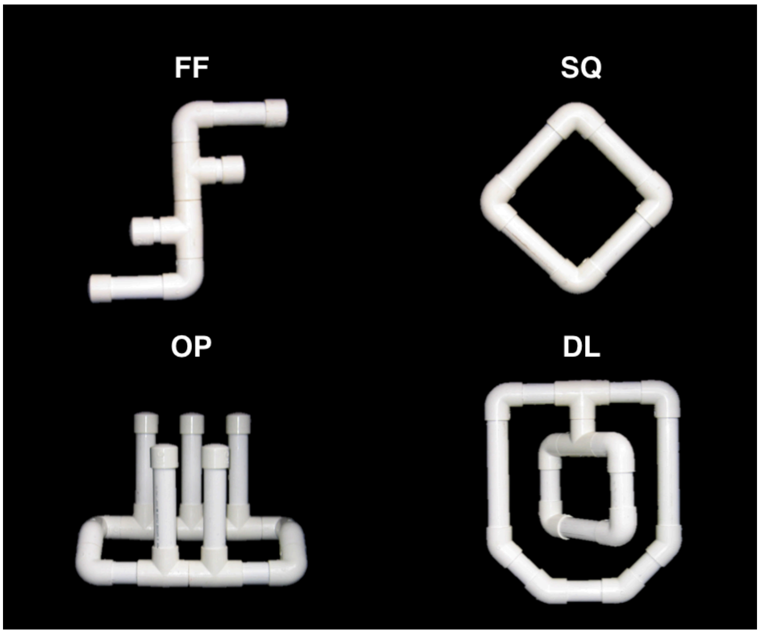


Figure 4: Baseline objects
FF and SQ were the baseline objects initially introduced followed by OP and DL.

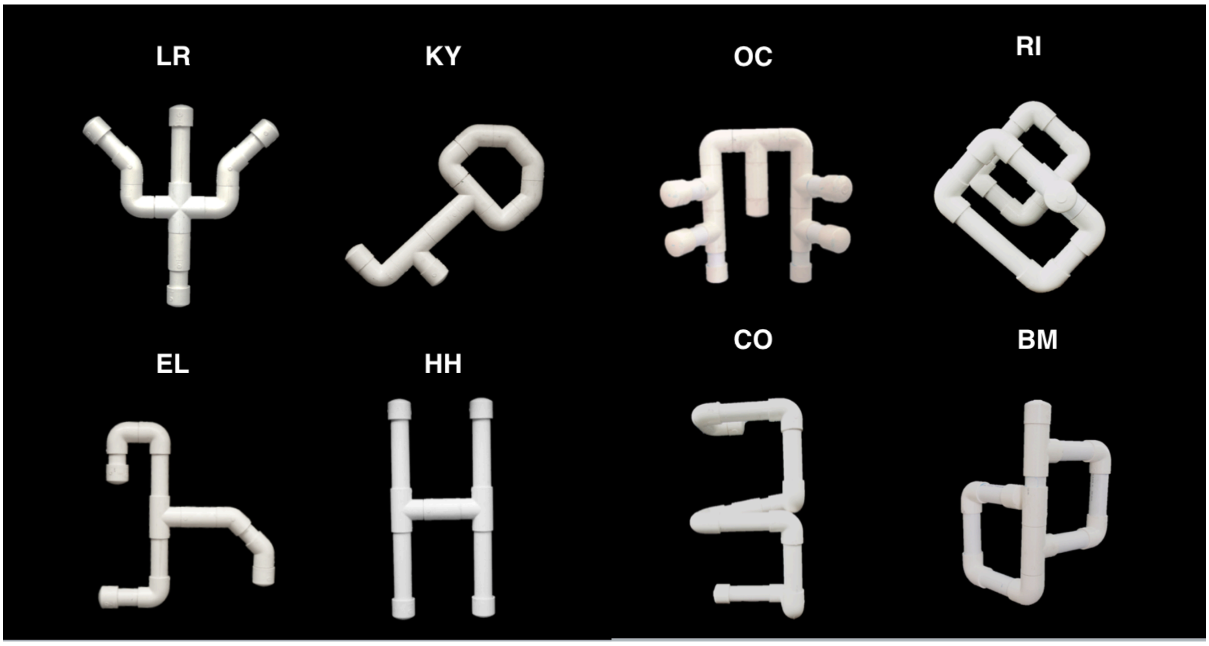


Figure 5: Set 1 and Set 2 objects
Set 1 (column 1 and 2) and Set2 (column 3 and 4) objects were introduced after all the baseline objects.

Methods

All the data analysis was performed using MATLAB.

The MTS training for the bottlenose dolphin Angelo started on the September 2, 2013. The data for the MTS training was analysed from the starting until the March 9, 2015. The data was analysed separately for visual-to-visual MTS task and echoic-to-visual MTS task. In order to plot the learning curves for the training, the data was further separated into sections across time. The starting of each section was determined by the introduction of new methods or objects for the training (Table 2). The data in each section was further divided according to the number of alternatives presented in the MTS tasks. Thus, the learning curves were plotted separately for a particular type of MTS task (visual-to-visual or echoic-to-visual) as well as for the number of alternatives presented.

Section Number	Date at the start of section (DD/MM/YYYY)	New training method at the start of the Section
Section 1	03/09/2013	Starting of MTS training
Section 2	17/01/2014	Baseline objects introduced for MTS V-V training
Section 3	18/02/2014	Set 1 and Set 2 objects introduced for MTS V-V training
Section 4	12/03/2014	MTS E-V training started
Section 5	16/05/2014	4 Alternative objects introduced for MTS V-V training
Section 6	26/09/2014	3 Alternative objects introduced for MTS V-V training
Section 7	15/12/2014	Set 1 and Set 2 objects introduced for MTS E-V training

Table 2: Sections of the training data for analysis
The data was sectioned into seven parts with respect to training time in order to analyse the responses for trials.

Percent ratio of correct responses (ROCR)

The parameter used to represent the learning curves describing Angelo's performance in executing the task, was the ratio of the correct responses (ROCR) calculated for all the trials conducted in a day, represented as percentage. The

ROCR was calculated as-

$$\text{ROCR} = \left[\frac{\text{Total number of correct responses in a day}}{\text{Total number of correct and incorrect responses in a day}}\right] \times 100$$

Acoustic Analysis

The hydrophone array used for the recording consisted of 32 Teledyne Reson TC-4013 hydrophones. The signals were amplified by a set of custom-built amplifiers which had a frequency range of 8 to 160 KHz. The signal from the amplifiers was acquired at a sampling frequency of 500 KHz (per channel) through a National Instruments data acquisition system. All the channels were synchronised. The acoustic data for each echoic-to-visual MTS trial was recorded on 16 hydrophones. The hydrophones were arranged on the sample box in a planar pattern (Figure 6)

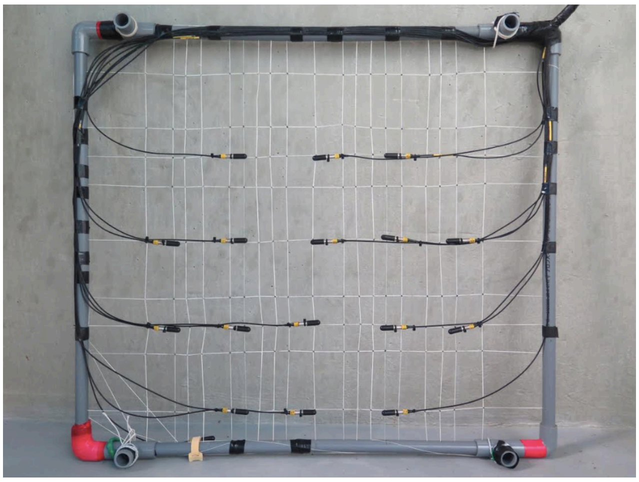


Figure 6: Hydrophone positions
The individual hydrophones are mounted on a frame with a mesh which is then placed on the anechoic box during the echoic-to-visual trials. (Between 2 hydrophones, horizontal distance - 12.5 cm, vertical distance - 15 cm)

The sample object was presented in the centre of the box. Thus, some hydrophones were in the proximity of the object as opposed to the rest. Out of these 16 hydrophones, data from 15 hydrophones had good recordings and were considered for the acoustic analysis. In some trials bubbles were used as an acoustic screen in an attempt to block the dolphin echolocation signal while the dolphin was still approaching the anechoic sample box and when he had left the echolocation box after echolocating. This gave a more precise time window of when

Angelo was actually able to use his echolocation sense to interrogate the sample, since it was nearly impossible for him to echolocate when the bubble screen was present. Thus, the data for each trial was time windowed to span only the period when the bubble screen was off.

1. Click Detection

The acoustic data for the echoic-to-visual MTS trial was initially analysed to detect the clicks. Among the 15 hydrophones that recorded the underwater audio, the one which recorded data that had the maximum total amplitude of signal recorded was used for the analysis. The clicks used for echolocating by a bottlenose dolphin generally lie below 150 kHz (Au, 1993). A bandpass filter in the frequency range of 30 to 160 kHz was used since the highest signal-to-noise ratio (SNR) of dolphin clicks lie within this range. The lower range of 30 kHz was set in order to filter out the low-frequency underwater noise. A Hilbert transform was applied on this bandpass-filtered data. The Hilbert transform yields the discrete time analytic signal based on the original real-valued signal, and its magnitude yields an envelope of the clicks, smoothening out some of the oscillations.

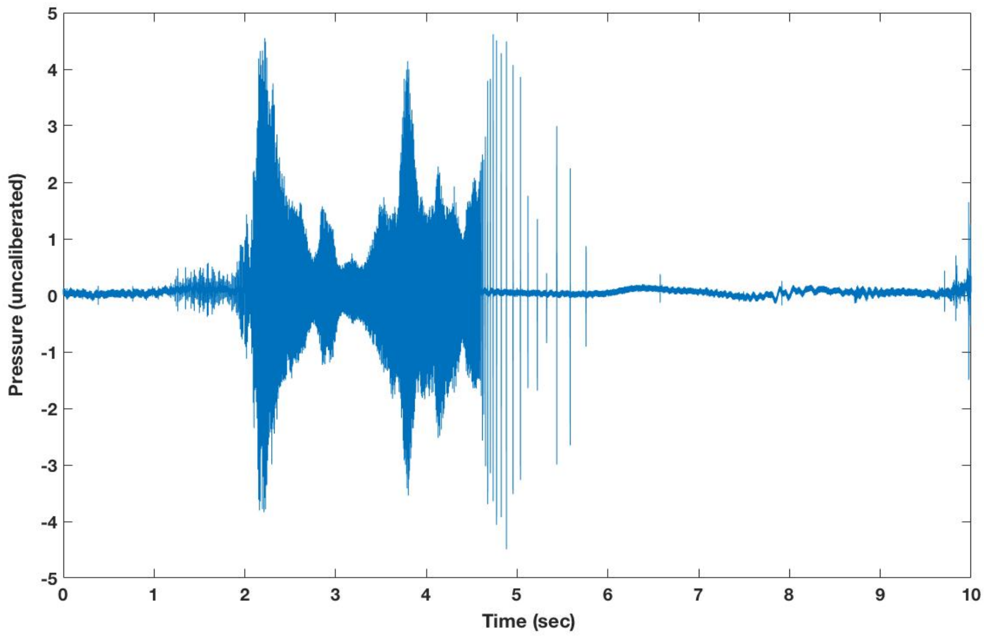


Figure 7: Original real-valued echolocation signal
In the figure most of the high amplitude signal is the representation of clicks. The click amplitude varies across the trial.

In each trial, the amplitudes of clicks used were not constant across time (Figure 7). Although the change in the amplitude was gradual, overall the amplitude varied a lot within the interrogation time duration, especially during the training phase. Using an adaptive threshold helped to detect both the low and high amplitude clicks. In order to make sure that noise was not detected as low amplitude clicks, a minimum threshold was set. Thus, using multiple thresholds helped to detect the clicks with a higher accuracy.

Double threshold for detecting clicks

Applying a bandpass filter on the data only partially eliminated the noise. The filtered data still had a certain noise component throughout. Thus, the first (lower valued) threshold was set using a part of the acoustic recording that did not have any dolphin echolocation signal. In order to detect such a part of the acoustic recording, the data was sectioned into either 400 or 500 equal parts (depending on the length of the selected data). Then the maximum value from a section with the lowest standard deviation was selected as the first threshold. The second threshold was determined using a moving window across six equally divided parts of the absolute values of the Hilbert transformed data. The Elbow method was used in order to set the second threshold. This method involved sorting the envelope of the recorded data in an ascending order, and finding a point on the distribution curve of the sorted data that had the shortest distance to the line joining the end points of the curve. This point was used as a reference to further select a higher threshold. This threshold was selected using a constant multiple (0.6) of the distance between the reference point and the end point on the curve. Thus, the moving window gave an adaptive threshold for each section of the data.

Along with the clicks, their reflections were recorded too. The six sections of the data envelope that were used to calculate the second threshold were subject to an additional step of envelope detection, using a window of fixed length (Figure 8). This window length for the Hilbert transform for this second round of envelope detection was determined by considering the length of a click and its possible reflection. Enveloping the data with the implementation of the second Hilbert transform helped to clump the transmitted clicks and their reflections (if present in the recording) as a single unit and not as two different clicks. In the envelope of the data, there could still be a few values higher than the first threshold that are not clicks but noise and correspond to the peaked envelopes. In order to select peaked envelopes that possibly corresponded to a click, the median of the values higher than the median of an envelope was used as a threshold. All the values that cross the first and the second threshold along with the threshold set with the envelope values were considered as clicks (Figure 9).

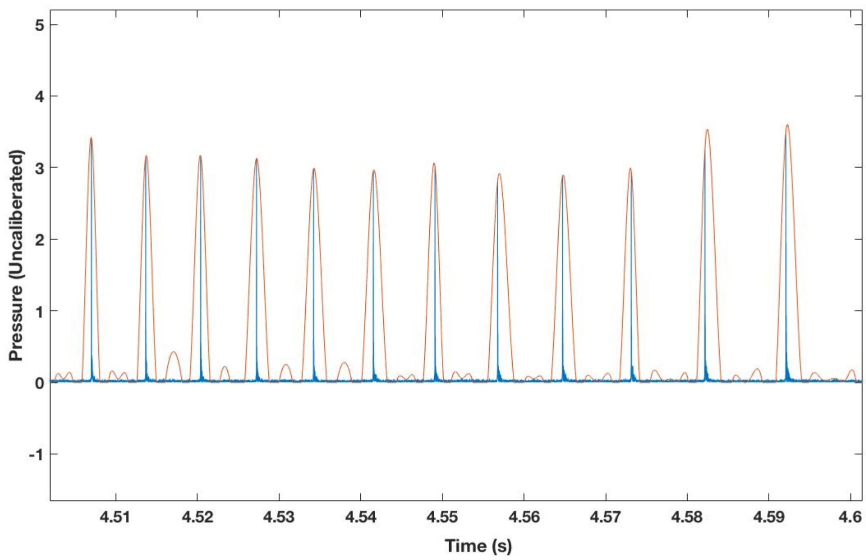


Figure 8: Representation of the first and the second Hilbert transform
The values plotted in blue represent the Hilbert transform of the original data. The values plotted in orange represent the second Hilbert transform.

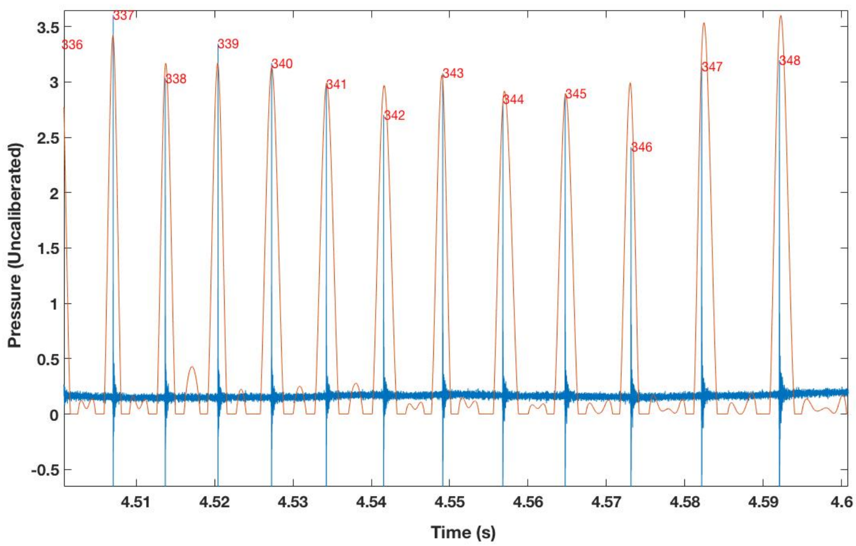


Figure 9: Detected clicks
The data represented in blue colour is the original real-valued data. The data represented in orange is the second Hilbert transform. The numbers from 336 to 348 correspond to the rank of the detected click.

Inter-click interval (ICI)

Inter-click interval was defined as the time interval between two consecutive clicks. The occurrence speed of the clicks transmitted depended on the time required for the reflection of the previous click to be received. If the target was far away, the time for the click's reflection to be received was long. Thus, the ICI would also be high as compared to when Angelo was near the object and echolocating on it. In the trials in which Angelo echolocated on the object for a time period shorter than the recording time, the clicks at the end which had a higher inter-click were deducted. All the files were manually screened to select the detected clicks in case of such trials.

2. Frequency Analysis

Along with the number of clicks, the bandwidth of the click signal used was another acoustic parameter that was analysed. In order to extract the spectral content of the clicks, for each click, a channel that recorded the highest amplitude was selected. When a dolphin echolocates, it emits a focused acoustic beam from its melon (an organ part of the forehead of a bottlenose dolphin). This beam is directed in a forward direction and thus the beam focus changes as the dolphins moves (Cranford, 2000). The hydrophones that were exactly in the direction of the beam recorded the clicks at an amplitude higher than any other hydrophone. Also, the high frequency components of the clicks are emitted in narrow mainlobe width while the low frequency components are in a wide mainlobe width (Au et al., 1986). Thus, in order to get an accurate description of the spectral content of the clicks, it was necessary to consider the direction of the echolocation beam. Computing spectral characteristics from the channel that gave the highest amplitude data for every click, yielded a more reliable estimate than selecting them from a single channel consistently.

In order to select the channel with the highest amplitude of each click, initially selected clicks (selection based on the ICI) from one reference channel were considered. A window length was computed for the possible recording of the click in other channels depending on the speed of sound underwater and the relative distance between the hydrophones. Thus, all the remaining channels were analysed for the highest amplitude recording of each click. A constant length of 256 data points or 512 μs was considered as a click. This data was considered asymmetrically across the highest amplitude point since clicks extended slightly more in the forward direction on the time series. The data for each click was bandpass filtered in the range of 30 to 160 kHz. A Hamming-window was applied to the data sets of all the clicks before calculating the Fast Fourier Transform. The power of the click signal corresponding to each frequency bin from 30 to 160 kHz was then computed. The power spectral density plots were normalised with respect to the maximum value for each click.

We defined an acoustic parameter named 'mean frequency' to collectively capture the spectral variation of click power across frequencies. The mean frequency indicated the frequency around which most of the acoustic energy was present. The mean frequency was calculated for each click as the mean of the power spectral density values represented in decibels. The decibel values were first normalised in such a way that the minimum was zero. The mean frequency was then computed using the following formula:

$$\text{Mean Frequency of click} = \frac{\sum(\text{Power spectral density} \times \text{Frequency})}{\sum \text{Power spectral density}}$$

Since the mean frequency is the weighted mean of the decibel values, in essence it represents the geometric mean of the click power distribution across frequency, which is also a good estimate of the mode of the spectral distribution (Gonzalez et al., 2006)

In order to combine the mean frequencies for all the clicks into a single simplified parameter that captured the spectral characteristics of clicks employed for each trial, the mean of the mean frequencies and the variance of the mean frequencies of the selected clicks in the trials were computed. Spectrograms were plotted for the trials, and the range of the frequencies represented corresponded to the frequency range of the bandpass filter used at the beginning of the frequency analysis. The power spectral density in the spectrograms was scaled in order to get an appropriate representation of the data.

For both the visual-to-visual and the echoic-to-visual trials, video recordings were available. A qualitative analysis of the video clips was performed whenever required in order to select the correct duration of data for the acoustic analysis.

Results and Discussion

1. ROCR/ learning curves

In behavioural studies involving training of animals, learning curves prove to be essential to understand the rate of learning over time (Harlow, 1949). For the matching-to-sample tasks, the ratio of correct responses (ROCR) was plotted across the complete duration of training. The ROCR for the trials across a certain number of days (averaged with an overlap across certain number of days on which trials were run) was calculated. The window for averaging the data was symmetric across the day being considered (-n/2 to +n/2 days, where n is the number of days over which the averaging is done). The data was analysed separately for each section as well as for each different trial. Echoic-to-visual and visual-to-visual were distinct tasks, therefore, their ROCR values were plotted separately. Thus, the computed ROCRs were plotted for six types of trials in total - visual-to-visual and echoic-to-visual tasks involving 2,3 or 4 alternatives (ALT). The confidence intervals for each ROCR value were calculated considering a binomial distribution of the task performance data and then using the Clopper-Pearson method (with 5% significance) (Johnson et al., 2005). Larger error bars represented more uncertainty in the estimated ROCR stemming from availability of fewer data points in the averaging.

In sections 1 to 3, there were only visual-to-visual MTS trials conducted. Since the training started with simple trials in which Angelo was clued to pick the correct alternative, the learning curve started with higher ROCR values. As the training progressed, the ROCR values eventually decreased. This could be explained by the new training methods (e.g., introduction of boxes and blinds to present objects, changes in the position of the boxes, changes in the object presentation methods) that were introduced during this phase of the training, thus increasing the probability of wrong responses as Angelo was in the process of understanding the task. Near to the end of section 1, the ROCR values started to decrease again. Information was available for whenever Angelo was hesitant to pick an object from the alternatives during the trials. This hesitant behaviour was defined as inconsistent behaviour. The inconsistent behaviour of Angelo near to the end of section 1 could have affected his responses. Since PVC objects were newly introduced in the training, the ROCR values decreased at the beginning of the section 2. The curve further increased as the section progressed, but again showed a drop at the beginning of section 3 which corresponded to the introduction of new PVC objects. The ROCR values remained constant in section 3. Overall the ROCR curve in sections 1 to 3 represented the understanding of the visual-to-visual task by Angelo (Figure. 10).

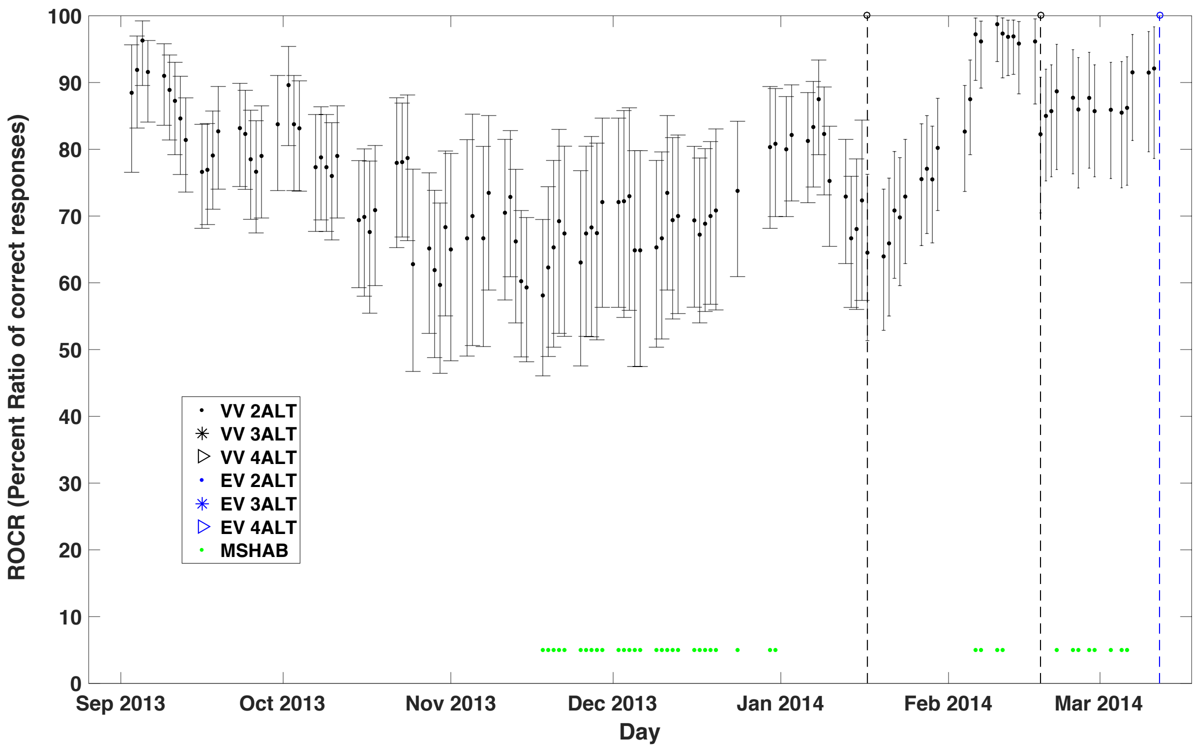


Figure 10: ROCR analysis of section 1, 2 and 3
The ROCR data averaged across four days for visual-to-visual MTS trials has been plotted. The error bars represent a 95% confidence interval. Each vertical line in the plot corresponds to the end of sections 1 to 3 respectively. MSHAB corresponds to the habituation sessions (represented as constant arbitrary values on the ROCR axis).

The beginning of the section 4 corresponded to introduction of the echoic-to-visual tasks in the training. The ROCR values for the echoic-to-visual tasks were lower than the ROCR values for the visual-to-visual tasks in the previous section. However, in the same section the visual-to visual matching tasks had values comparable to that in the previous section, thus suggesting that Angelo had acquired an understanding of the visual-to-visual task by then. In the later part of the section 4, the ROCR values for the echoic-to-visual tasks dropped to almost half as compared to those at the beginning. However, the large error bars for these values also suggested that the low ROCR values could be due to the fact that the available data points used for its calculation were few in number. Four-alternative visual-to-visual tasks were introduced at the beginning of the section 5 but only for a few trials. The corresponding ROCR values dropped, though there were only a few trials indicating this, as shown by the large error bars. There was a large gap in the training during the session 5. This resulted in the decrease in the ROCR values later. However, a better understanding of the visual-to-visual task led to an immediate rise in their ROCR values again (Figure 11).

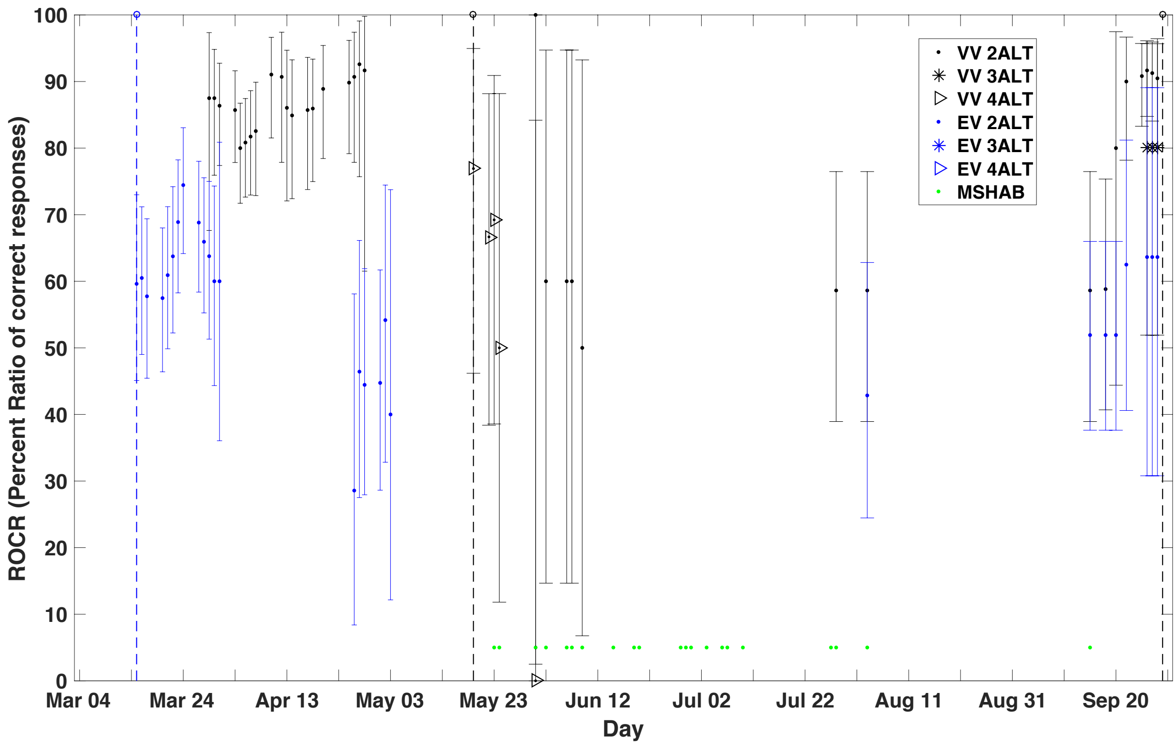


Figure 11: ROCR analysis of section 4 and 5

The ROCR data averaged across four days for different trial has been plotted for the months in the year 2014. The error bars represent a 95% confidence interval. The first two vertical lines in the plot correspond to the beginning of section 4 and 5 respectively. MSHAB corresponds to the habituation sessions (represented as constant arbitrary values on the ROCR axis).

In section 6, the visual-to-visual trials began with the introduction of the three-alternative tasks (Figure 12). The four-alternative tasks were introduced in this section as well. However, the ROCR values were fairly high throughout both the section 6 and 7. Thus, introduction of a new object has a higher effect on the responses than the change (increase) in the number of alternatives presented. This was indicative of the fact that responses for the visual-to-visual trials were not merely due to chance but a result of a conscious choice made by Angelo.

Section 6 and 7 had the highest number of echoic-to-visual matching tasks. In section 6 as well, Angelo had still not completely understood the echoic-to-visual matching task. The learning curve of the visual-to-visual matching tasks was always higher than that for the echoic-to-visual matching tasks. The increase and drop in the ROCR values of the echoic-to-visual tasks in November 2014 exactly corresponded to the start and end of the training with the partially covered Plexiglas. However, the ROCR values dropped even further after the end of the training with the partially

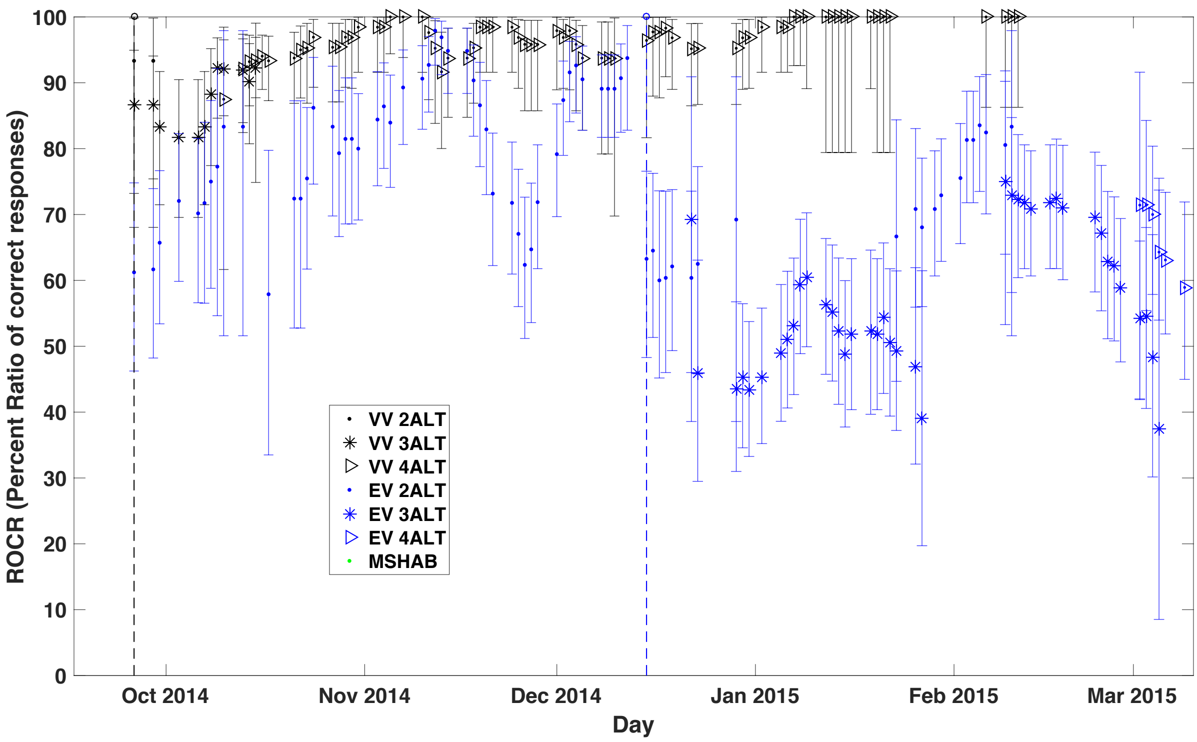


Figure 12: ROCR analysis for section 6 and 7

The ROCR data averaged across four days for different trial has been plotted. The error bars represent a 95% confidence interval. The two vertical lines in the plot correspond to the beginning of the section 6 and 7 respectively.

covered Plexiglas training on November 13, 2014, but showed an increase or ‘recovery’ later. This was explained by two possible reasons. Firstly, since the sample box was partially covered, Angelo could have still used whatever visual cues were available in order to perform the MTS task, while his better understanding made him more proficient at the task. Alternatively, even if the set up were new, Angelo may have eventually begun translating his existing understanding of a visual-to-visual task into an echoic interrogation, and thus could perform better. Either of these explained the gradual increase in the ROCR values even if the Plexiglas cover increased during this training.

During the training with the partially covered Plexiglas, the object was presented both in the visible part of the box and behind the Plexiglas. Hence along with the visual perception of the objects, Angelo must have also tried to echolocate the object in these trials. Thus, when this training ended, he was still able to perform the echoic-to-visual task well. FF and SQ were the objects used during the training with the partially covered Plexiglas. They were the only objects used for a few trials after this training as well. Later, another set of baseline objects was introduced (OP and DL); the drop in the ROCR values after the training corresponds to this event. However, eventually the values increased again. Since these increased ROCR values for the echoic-to-visual trials were higher than before the training with the partially covered Plexiglas started, it can be concluded that the training was successful.

The starting of the section 7 corresponds to the introduction of the Set 1 and Set 2 objects for the echoic-to-visual trials. This explains the drop in the ROCR curve at the starting of section 7 for trials even with two alternatives. The section 7 also had three-alternative and four-alternative echoic-to-visual trials. The ROCR values of the two-alternative trials showed an increase, indicating a better understanding of the two-alternative echoic-to-visual task even with multiple objects (including Baseline, Set 1 and Set 2 objects - 12 in total). The ROCR values for the three-alternative and the four-alternative echoic-to-visual tasks with multiple objects remained more or less constant and below 80%. Angelo’s performance for these three-alternative and four-alternative tasks was above 33.3% and 25% respectively which were the probabilities of getting the match correct by random chance; this suggested an understanding of the task by Angelo. Additionally, there was a slight drop in the ROCR values for these trials at the end of the section which corresponded to a more frequent use of the Set 1 and Set 2 objects than before.

Overall the trends in the ROCR values for both the visual-to-visual and echoic-to-visual tasks in all the sections suggest an improved understanding of the MTS task even with the introduction of every new object or training method.

2. Acoustic parameters

a.) Numbers of Clicks

The initial acoustic parameter considered for the analysis was the number of clicks used in the echoic-to-visual MTS tasks. The number of clicks averaged over a certain number of days for trials with a particular number of alternatives was compared with an average ROCR value calculated for the corresponding trials. In order to get the confidence intervals for the average number of clicks, the hypothesis that the set of click numbers considered had a Gaussian distribution was tested using the Kolmogorov-Smirnov test (K-S test) (Massey, 1951). The test was applied on the number of clicks corresponding to the trials that were considered to plot a value for the average click number per day. Since the data was averaged across four days, the number of clicks for the trials across four days was considered as one data set. The K-S test tested the goodness of fit of the cumulative distribution of the empirical data set with that of the cumulative distribution function of a standard Gaussian distribution with a 5% significance. A significant number of click number data sets passed the K-S test and thus the confidence intervals for the average click number were computed under the assumption that it followed a Gaussian distribution (Figure 13).

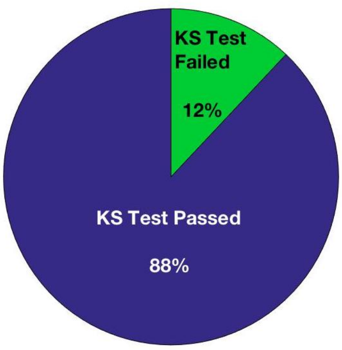


Figure 13: Kolmogorov-Smirnov test (K-S test) results
The diagram shows a higher number of click data sets that have passed the K-S test as compared to the data sets that failed the test.

The comparison of the ROCR with the average click number showed that the distribution of both ROCR and average click number was independent of each other for most of the days (Figure 14). Particularly, in the section 7, the number of clicks were almost constant throughout the section irrespective of the type of trials being considered. In the section 6, the number of clicks increase at the point of time when the training with the partially covered Plexiglas began. However, this increase could be because of the methods involved in the training that increased the target localisation time. Additionally, in section 4, the echoic-to-visual trials were newly introduced thus leading to Angelo localising on the target for a longer time. This led to increased number of clicks.

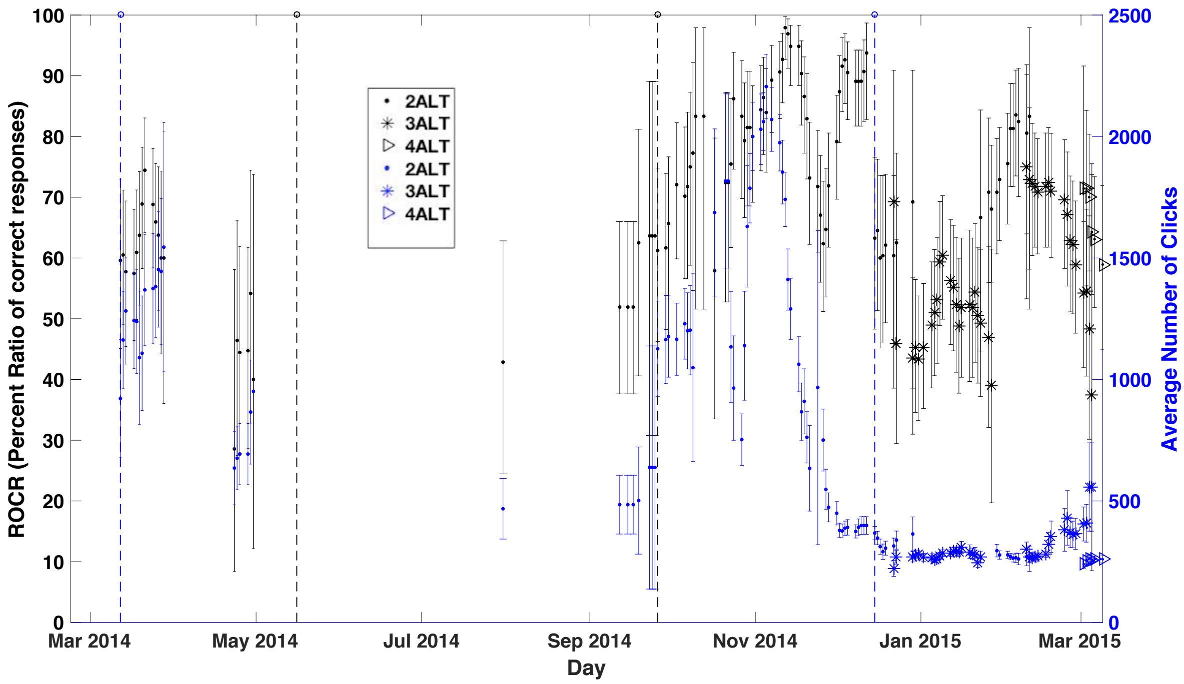


Figure 14: ROCR and Click number analysis

The ROCR and click number data averaged across four days for different trial has been plotted. The error bars represent a 95% confidence interval. The four vertical lines in the plot correspond to the beginning of the sections 4, 5, 6, and 7, respectively. The values plotted in black represent the ROCR values while the ones in blue represent the average number of clicks.

An N-way ANOVA was performed to check the effect of the click number (averaged per day), number of alternatives and the point of time of training on the ROCR values (averaged per day). The click numbers were grouped as low, medium and high depending on the value. While the time was considered in the groups of months (11 months). The ROCR was not significantly different for the number of clicks (low, medium, high) (p-value = 0.75) or for the number of alternatives (2, 3, 4) (p-value = 0.14). However the ROCR varied significantly (p-value < 0.05) with the time of training.

The number of clicks for a trial might be different depending on the sample object being presented. Thus, the data was analysed separately for all the 12 objects used in the trials. The median number of clicks were considered for all the different objects (median was selected as the parameter since it would be unaffected by the outliers present in case of a few objects). The median number of clicks were observed to be high for the baseline objects (FF and SQ) that were initially introduced for the echoic-to-visual trials (Figure 15). These were also the only objects used during the training with the partially covered Plexiglas. For the other baseline objects (OP and DL), the median values were not as high as for FF and SQ but they were still greater than those for the Set 1 and Set 2 objects. The median click numbers for the Set 1 and Set 2 objects were in a very close range. Thus, the click number was greater if the object was introduced earlier in the training.

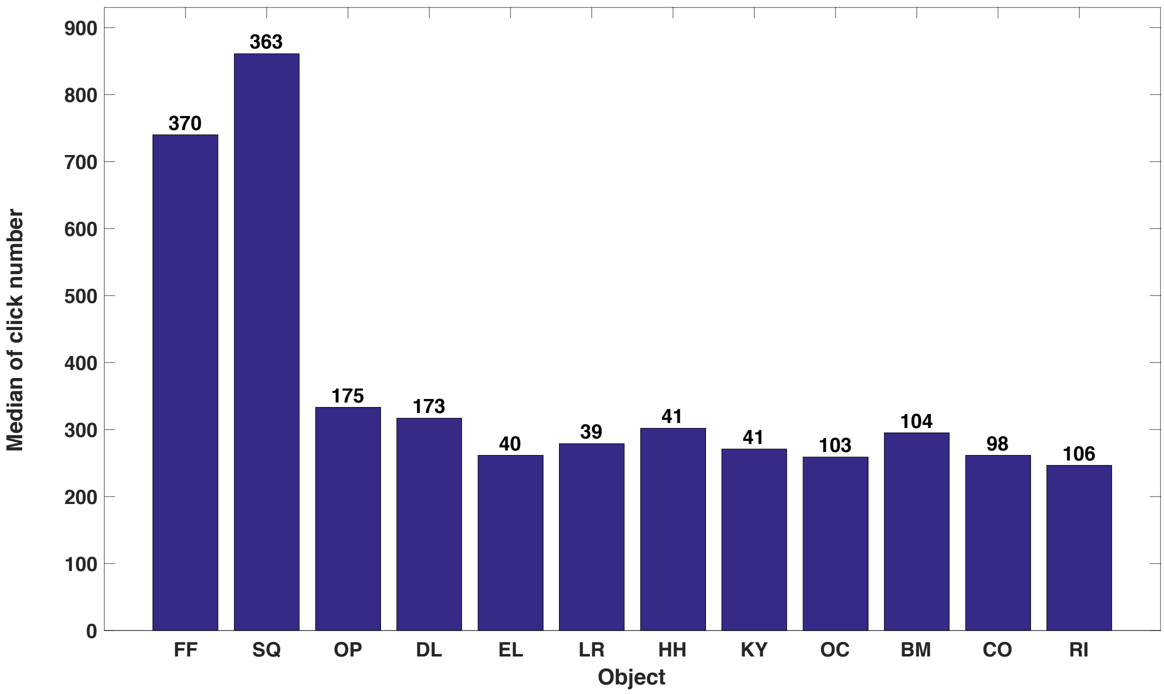


Figure 15: Median number of clicks for all the objects
The numbers on top of the bars represent the total number of trials for each object.

The probability density of the number of clicks gave an estimate of the distribution of the clicks used in case of each object (Figure 16). The number of clicks used by Angelo in the trials was comparable for many of the objects. The number of clicks

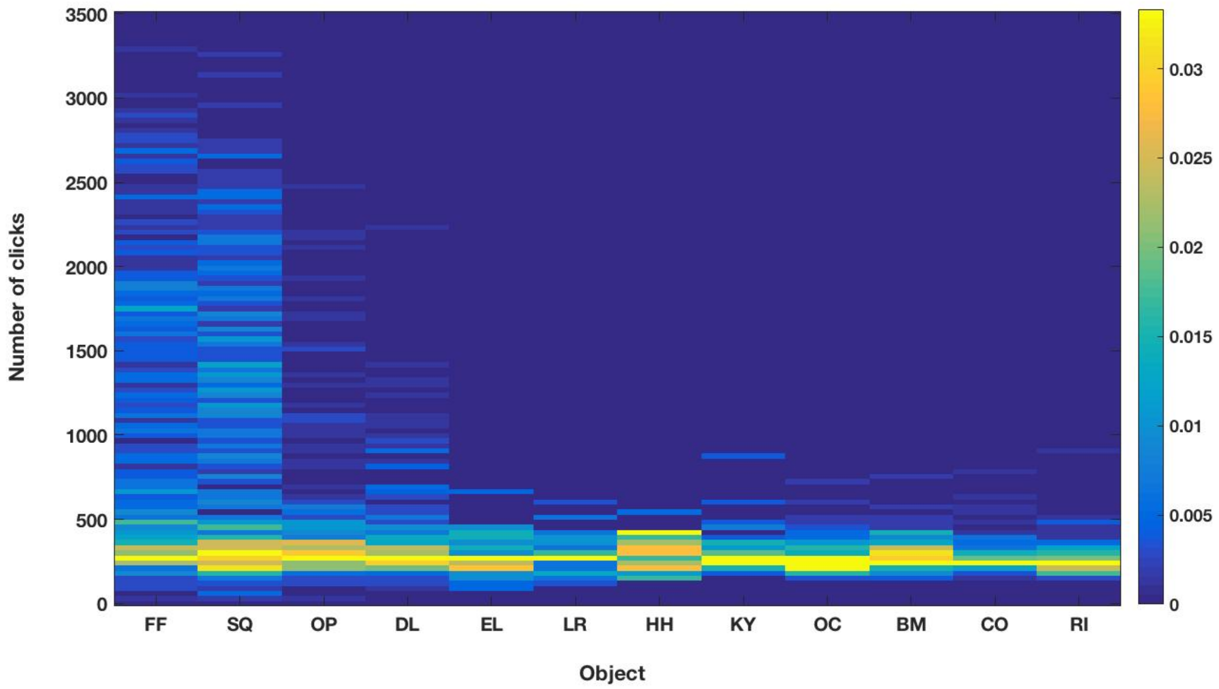


Figure 16: Probability density of clicks per trial for all objects
The probability density of clicks was plotted per trial (normalised for trials) for all the objects.

used for SQ and FF were relatively higher, which could be because they were used in the initial part of the training. The number of clicks used for HH and BM were also slightly higher than the average for other objects. This could be because Angelo could have perceived these as complex objects thus making use of more clicks in the corresponding trials for these objects. The distribution of the number of clicks for trials with correct and incorrect outcomes for all the objects was compared (Figure 17). The distribution of the number of clicks for both type of trials was significantly different (Wilcoxon rank sum test: p = 7.1201e-08, 5% significance). Thus, the number of clicks could determine or correlated with a correct or incorrect response.

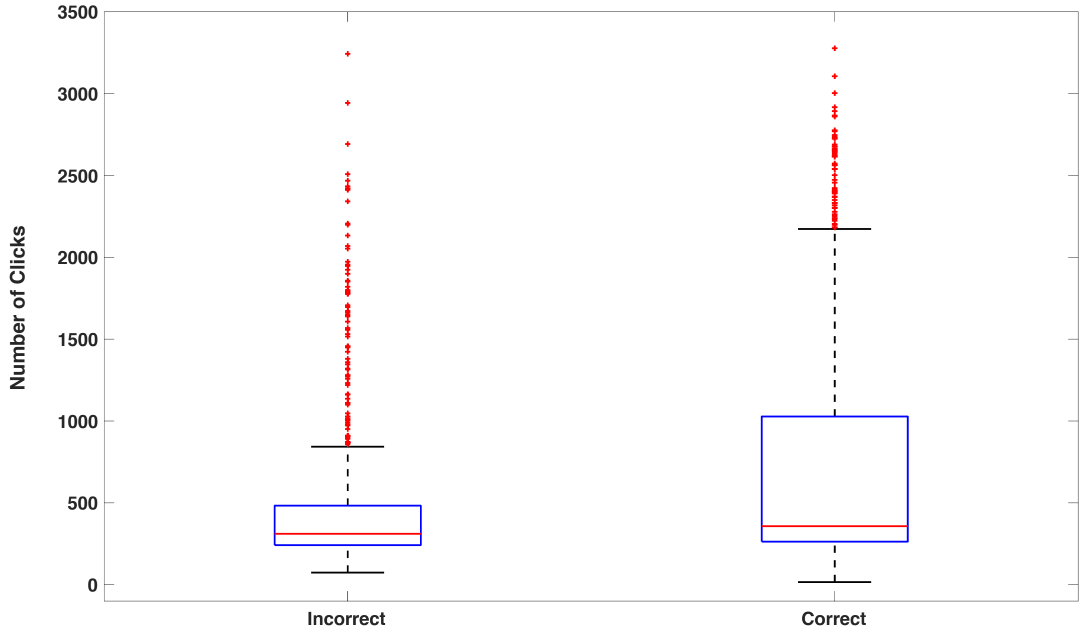


Figure 17: Distribution of number of clicks for correct and incorrect trials
The distribution of number of clicks was plotted for correct and incorrect trials in case of all the objects. The red lines in the boxes represent the medians for the respective group. The upper and lower boundaries of the boxes in the plot are the 75% and 25% quantiles of the data respectively. The markers in red represent the outliers of the data sets. The median values are significantly different with p = 7.1201e-08 (5% significance).

Although the average number of clicks plotted across the complete duration of training had a decreasing trend, there were some differences among the click number for individual objects. In case of the baseline objects FF and SQ, the trend for the number of clicks was very similar. However, for another baseline object pair, OP and DL, the number of clicks in the section from March to May 2014, were lower in case of DL as compared to OP (Figure 18). In the later part of the training, between November 2014 and January 2015, objects OP and DL were introduced again after a long gap and the number of clicks used for them was high as compared to that for FF and SQ which was lower around that point of time. But the trend was similar for OP and DL in the later part of the training. Similarly, in case of the Set 1 objects, except for a few outliers all the number of click values for all the objects were less than 600 (Figure 19 and 20). In case of the Set 2 trials also, except for a few outliers, the number of click values for all the objects were less than 600.

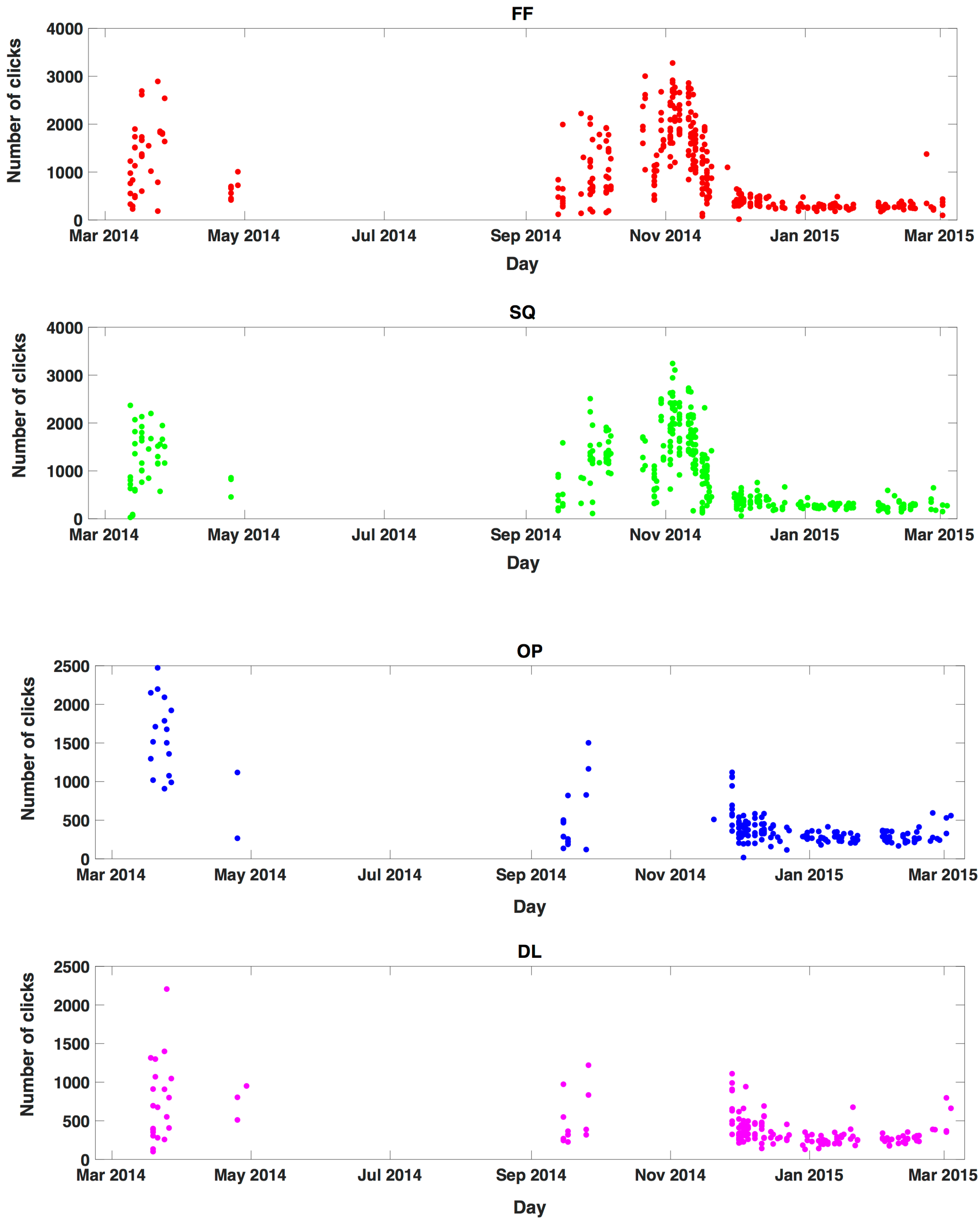


Figure 18: Number of clicks for all the trials for Baseline objects

The number of clicks for all the baseline objects has been plotted for all the individual trials across the training time.

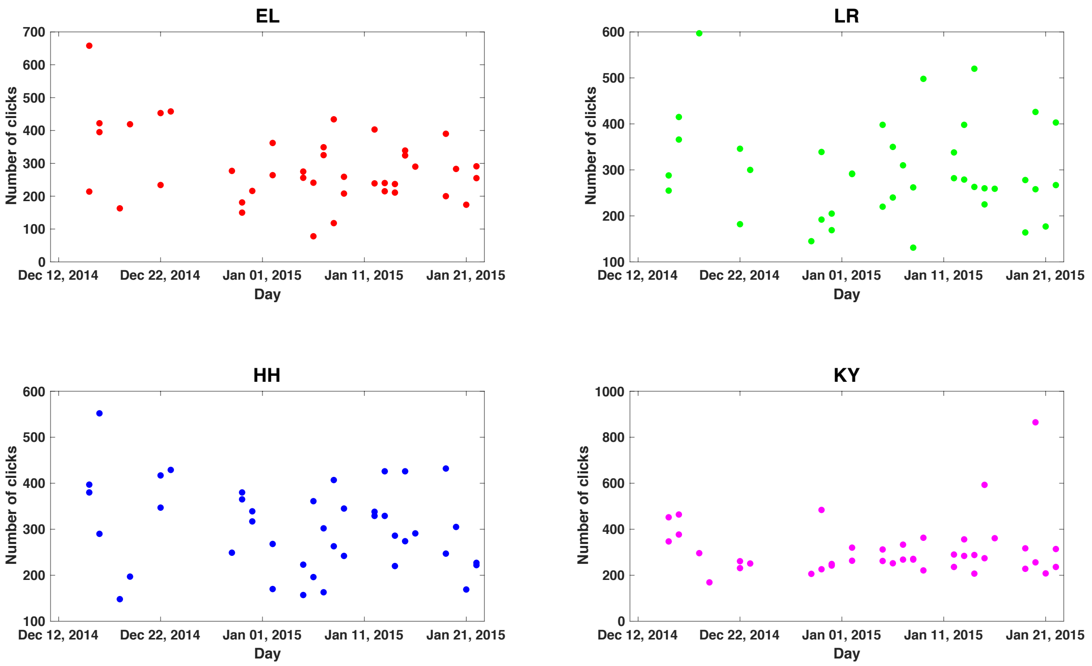


Figure 19: Number of clicks for all the trials for Set 1 objects

The number of clicks for all the Set 1 objects has been plotted for all the individual trials across the training time.

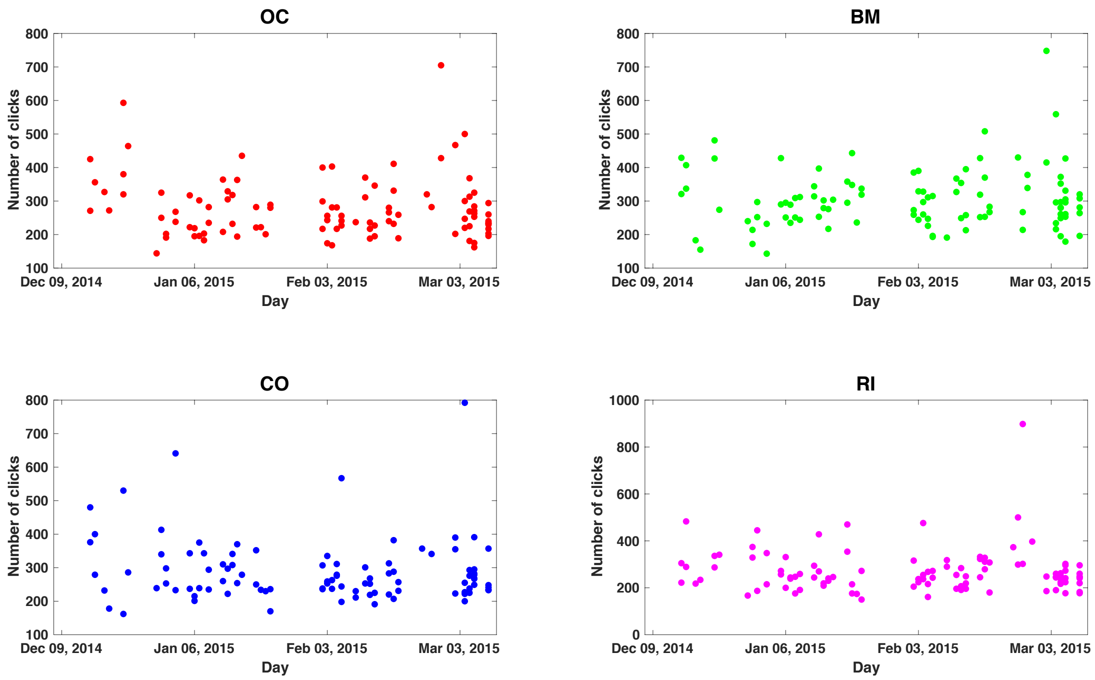


Figure 20: Number of clicks for all the trials for Set 2 objects

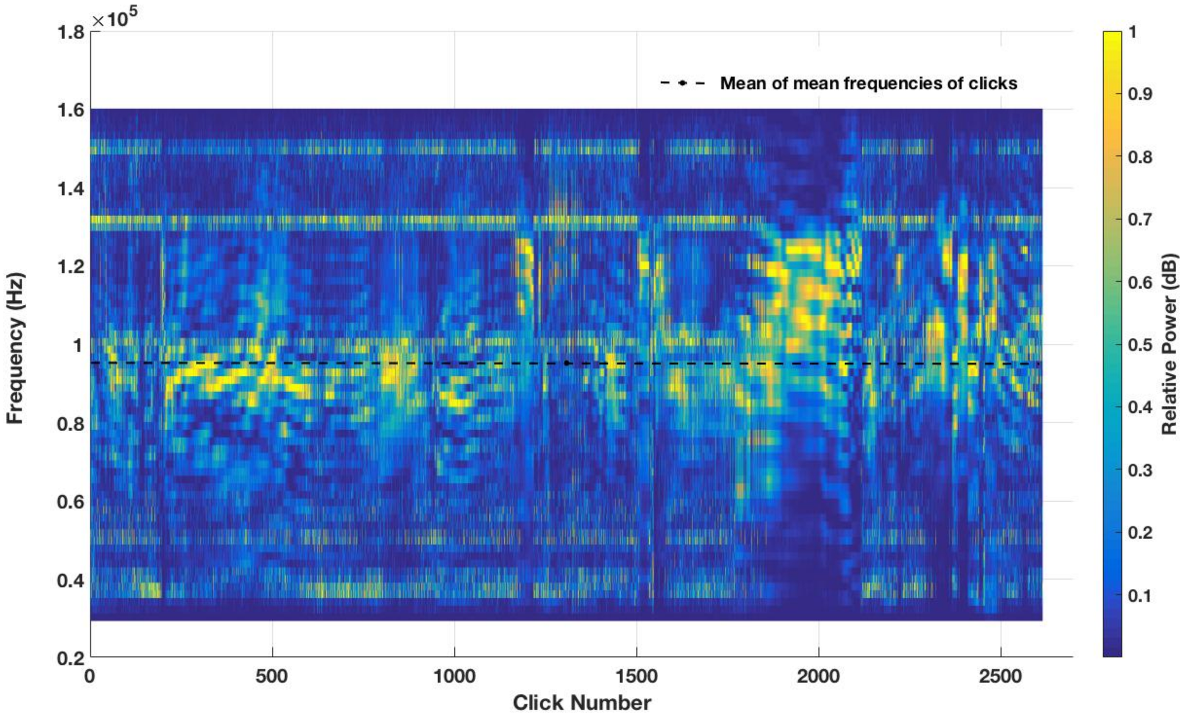
The number of clicks for all the Set 2 objects has been plotted for all the individual trials across the training time.

In the duration of training for the baseline objects corresponding to the complete duration of training for the Set 1 and Set 2 objects, the numbers of clicks mostly were below 600. Thus, the constant trend for the averaged number of clicks for all objects in the section 7 of the training (as plotted in Figure 11) reflects across all the objects and is not attributed to one object alone.

b.) Frequency Analysis

The spectrograms for the trials with FF and SQ used as sample objects, conducted in the initial phase of the echoic-to-visual training (March 2014) and in the last phase (March 2015) showed clear differences (Figure 21 and Figure 22). The power spectral density was plotted across clicks, not time. The power spectral density was not continuous if the clicks were not distributed uniformly across the time series; thus showing abrupt jumps in the data where there was a large time gap between clicks.

The mean of the mean frequencies of all the clicks in a trial was considered as one of the single parameters for the complete trial since it would give an estimate of the central tendency of the mean frequency in a trial. Although the mean of the mean frequencies of the clicks in the trials was close to 100 kHz for all the trials, the variance of the mean frequencies of the clicks was different for trials in each phase of training. Additionally, the power of the clicks was higher for trials in the later phase of training, as compared to the trials in the initial phase of training.



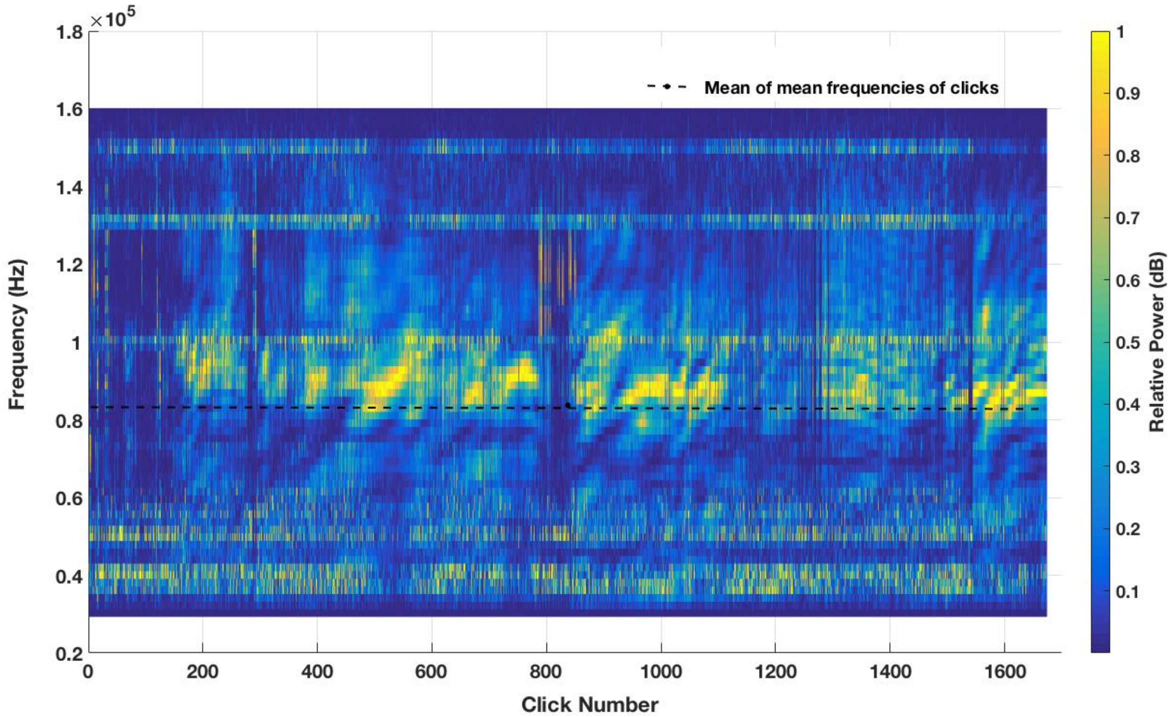
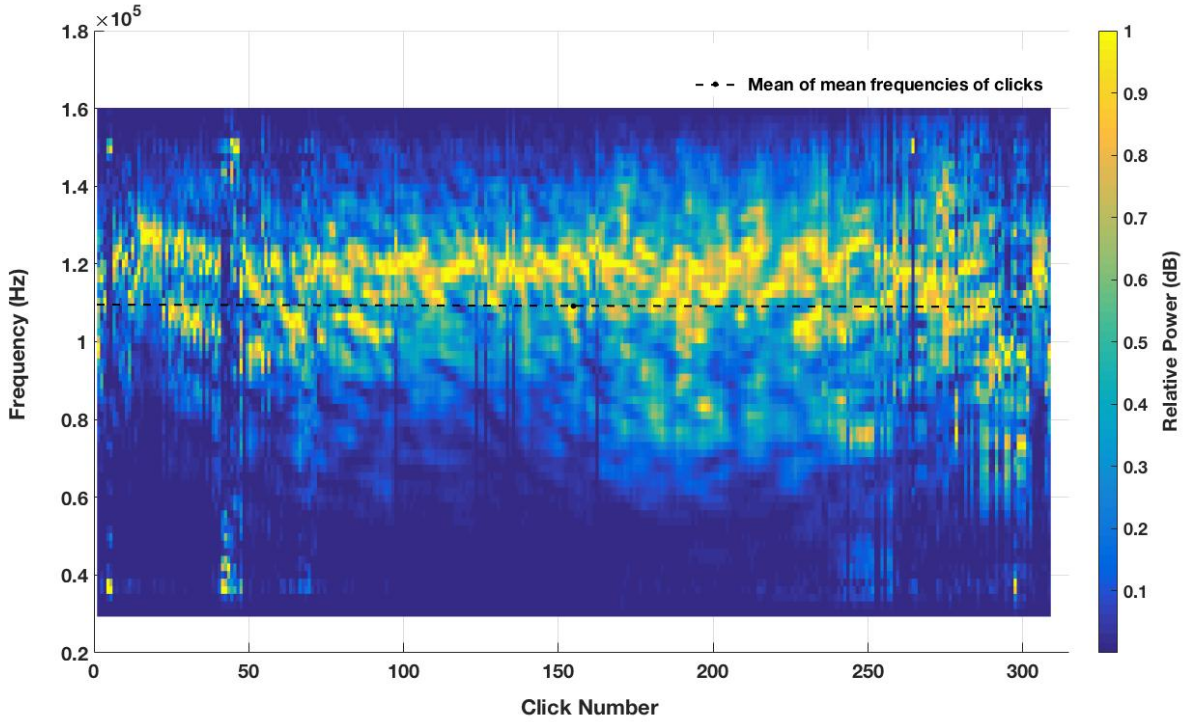


Figure 21: Spectrograms for trials conducted in the initial phase
The first spectrograms shows a trial conducted using FF as the sample object on March 17, 2014.
The second spectrogram shows a trial conducted using SQ as the sample object on March 21, 2014.



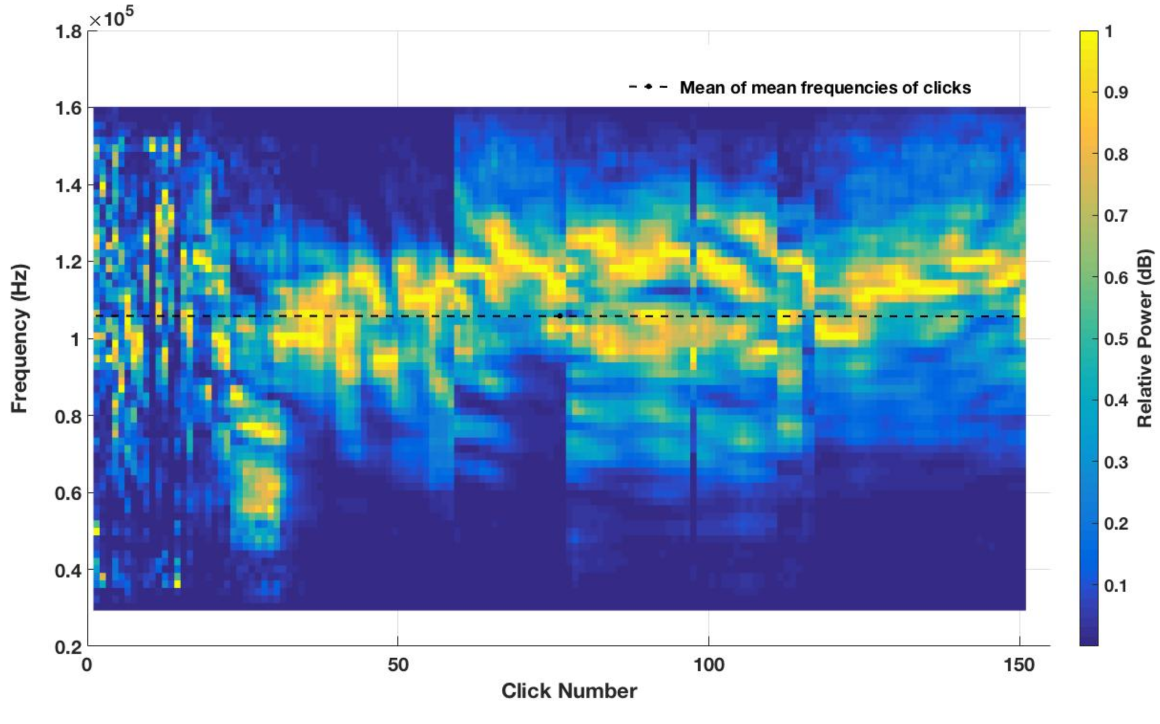
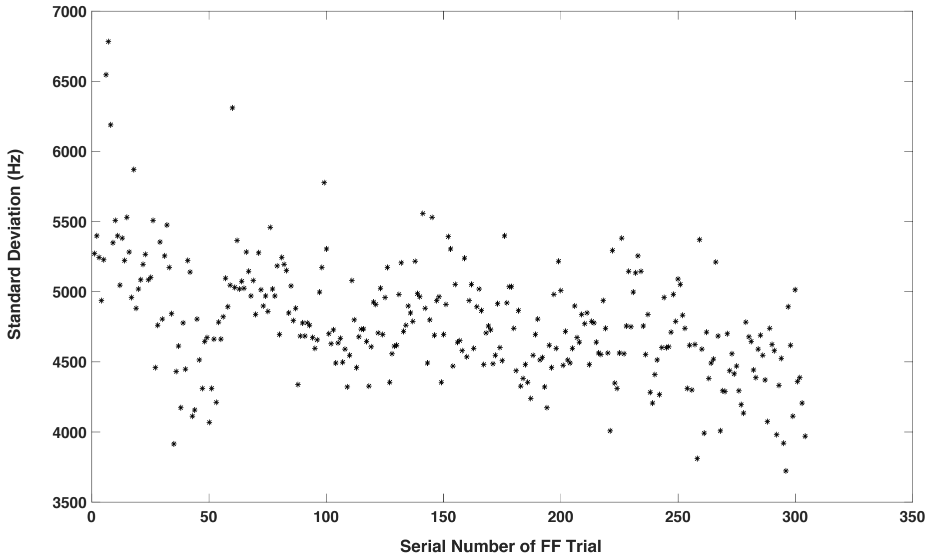


Figure 22: Spectrograms for trials conducted in the later phase
The first spectrograms shows a trial conducted using FF as the sample object on March 02, 2015.
The second spectrogram shows a trial conducted using SQ as the sample object on March 02, 2015.

The standard deviation of the mean of mean frequencies of clicks for the baseline object FF showed an overall decreasing trend across individual trials. Similarly, for the standard deviation of mean frequencies for the object SQ, there was a slight decrease after some initial trials, however, the values stayed constant later for many trials. Trials after 250 showed decreased standard deviation values. However, there were also a few that had values as high as in the initial trial (Figure 23).



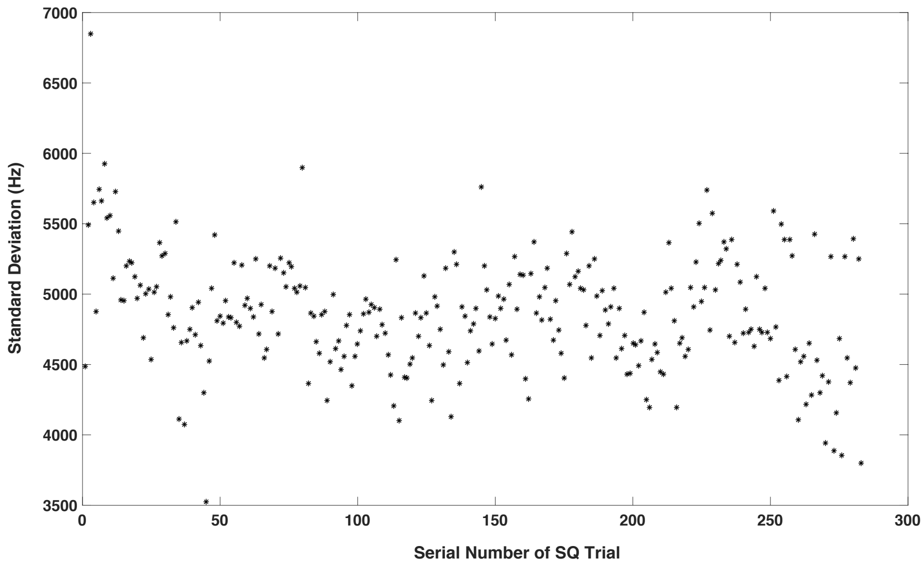


Figure 23: Standard deviation values of the mean frequencies of clicks
The first plot shows standard deviation values of the mean frequencies of clicks for all trials with FF as the sample object. The linear correlation coefficient (r) for these mean frequency values is - 0.4716. The second plot shows standard deviation values of the mean frequencies of clicks for all trials with SQ as the sample object. The linear correlation coefficient (r) for these mean frequency values is - 0.2186.

Summary

The ROCR values across the complete duration of the training showed trends that could be explained by the increased number of trials and the introduction of new objects and techniques. Angelo had understood the visual-to-visual trials very well and performed correctly almost every time in the last section of the training. For the echoic-to-visual task as well, even if the ROCR values were not very high in the last section, they were still consistently higher than chance values. Thus, there was an understanding of the echoic-to-visual tasks by Angelo. Introduction of new objects for the training had a significant effect on the ROCR values. However, the number of clicks were high only when the echoic-to-visual trials began or when there was a large gap in the training. Correspondingly, the number of clicks were high in case of the baseline objects that were introduced initially for the echoic-to-visual training. The number of clicks for the set 1 and set 2 objects that were introduced later were almost constant across trials. The number of clicks for two of the initially introduced baseline objects at the beginning of the training was slightly different. Additionally, the distribution of the number of clicks per trial was only slightly different for four objects, two of which were baseline objects. The distribution of the number of clicks for the correct and incorrect trials was significantly different. However, the number of clicks did not have a significant effect on the ROCR mean values, though the time of training did. Thus, the overall changes in the number of clicks suggest that once the task was understood, the click number did not play a significant role in the MTS task. However, depending on the dolphin's perception of the complexity of an object, the number of clicks could vary only for the initial echoic shape estimation of that object.

In case of the spectral characteristics of the clicks used for two of the initially introduced baseline objects, there were notable differences in the total power used and the variance of the mean frequencies of the clicks for the trials in the initial and the later phases of the training. By using a wider bandwidth of frequencies in the initial phase, Angelo would had been trying to set a range of frequencies to echoically estimate the object shape. He would had eventually adjusted the bandwidth to a narrower range to get the maximum information about the object through the echoes in order to accurately distinguish the object. Once the range of the frequencies was adjusted, Angelo would had been able to perform the task quickly, thus decreasing the number of clicks as well. Considering the lower number of clicks used when the set 1 and set 2 objects were introduced, there could have been some common factor in the baseline objects and the set 1 and set 2 objects that would have helped him estimate their echoic shape quicker. Overall, the bandwidth of frequencies used in the clicks played an important role in performing the MTS task.

The results show that there was fine-tuning of the acoustic parameters of the clicks by Angelo, also suggesting physical learning apart from cognitive understanding of the task which may have helped his performance. Further frequency analysis of the data can give better insights into the understanding of the echolocation behaviour. There could be acoustic factors other than the click numbers or the frequency bandwidth (e.g., inter-click interval) that would have changed during the training which could be analysed as well.

The development of a behaviour can involve both innate and learned aspects. In order to properly understand the innate aspect of the echolocation behaviour in dolphins, further experiments could be performed that record the neonate dolphins. Using devices that identify the vocalising animal could help record neonates even if they are in the presence of their mother or other dolphins (Hoffmann-Kuhnt et al., 2016). For this project the data was analysed for experiments that were performed in a captive set-up. This provided insights into the changes of the acoustic parameters across the matching-to-sample training. It would be interesting to compare these results to the changes in the acoustic parameters of a developing neonate in the wild.

References

Au, Whitlow WL, Patrick WB Moore, and Deborah Pawloski. "Echolocation transmitting beam of the Atlantic bottlenose dolphin." *The Journal of the Acoustical Society of America* 80, no. 2 (1986): 688-691.

Au, Whitlow WL. "Characteristics of dolphin sonar signals." In *The Sonar of Dolphins*, pp. 115-139. Springer, New York, NY, 1993.

Au, Whitlow WL. "Echolocation in dolphins." In *Hearing by whales and dolphins*, pp. 364-408. Springer, New York, NY, 2000.

Cranford, Ted W. "In search of impulse sound sources in odontocetes." In *Hearing by whales and dolphins*, pp. 109-155. Springer, New York, NY, 2000.

Gonzalez, Juan G., José L. Paredes, and Gonzalo R. Arce. "Zero-order statistics: a mathematical framework for the processing and characterization of very impulsive signals." *IEEE Transactions on Signal Processing* 54, no. 10 (2006): 3839-3851.

Gould, Edwin, Norman C. Negus, and Alvin Novick. "Evidence for echolocation in shrews." *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* 156, no. 1 (1964): 19-37.

Griffin, Donald R. "Measurements of the ultrasonic cries of bats." *The Journal of the Acoustical Society of America* 22, no. 2 (1950): 247-255.

Griffin, Donald R. "Acoustic orientation in the oil bird, Steatornis." *Proceedings of the National Academy of Sciences* 39, no. 8 (1953): 884-893.

Harder, Jennifer Hendry, Heather M. Hill, Kathleen M. Dudzinski, Karissa Tranel Sanabria, Sara Guarino, and Stan A. Kuczaj. "The Development of Echolocation in Bottlenose Dolphins." *International Journal of Comparative Psychology* 29, no. 1 (2016).

Harlow, Harry F. "The formation of learning sets." *Psychological review* 56, no. 1 (1949): 51.

Herman, Louis M., Michael F. Peacock, Michael P. Yunker, and Carolyn J. Madsen. "Bottle-nosed dolphin: double-slit pupil yields equivalent aerial and underwater diurnal acuity." *Science* 189, no. 4203 (1975): 650-652.

Herman, Louis M., Adam A. Pack, and Matthias Hoffmann-Kuhnt. "Seeing through sound: Dolphins (Tursiops truncatus) perceive the spatial structure of objects through echolocation." *Journal of Comparative Psychology* 112, no. 3 (1998): 292.

Hoffmann-Kuhnt, M., D. Herzing, A. Ho, and M. Chitre. "Whose line sound is it anyway? identifying the vocalizer on underwater video by localizing with a hydrophone array." *Animal Behavior and Cognition* 3, no. 4 (2016): 288-298.

Johnson, Norman L., Adrienne W. Kemp, and Samuel Kotz. *Univariate discrete distributions*. Vol. 444. John Wiley & Sons, 2005.

Massey Jr, Frank J. "The Kolmogorov-Smirnov test for goodness of fit." *Journal of the American statistical Association* 46, no. 253 (1951): 68-78.

Novick, Alvin. "Acoustic orientation in the cave swiftlet." *The Biological Bulletin* 117, no. 3 (1959): 497-503.

Pye, J. David. "A theory of echolocation by bats." *The Journal of Laryngology & Otology* 74, no. 10 (1960): 718-729.

Vroomen, Jean, and Beatrice de Gelder. "Sound enhances visual perception: cross-modal effects of auditory organization on vision." *Journal of experimental psychology: Human perception and performance* 26, no. 5 (2000): 1583.