Information-seeking across Auditory Scenes by an Echolocating Dolphin

Heidi E. Harley[1,2*], Wendi Fellner[2,3], Candice Frances[1], Amber Thomas[2], Barbara Losch[2], Katherine Newton[1], and David Feuerbach[2]

[1]Division of Social Sciences, New College of Florida, Sarasota, FL, United States

[2] The Seas, Epcot®, Walt Disney World® Resorts, Lake Buena Vista, FL, United States

[3]St. Andrews University, etc.

Author Note

Candice Frances is now at the Basque Center on Cognition, Brain and Language

Amber Thomas is now at The Pudding

Katherine Newton is now at…

Correspondence concerning this article should be addressed to Heidi E. Harley, Division of Social Sciences, New College of Florida, 5800 Bay Shore Road, Sarasota, FL, 34243, USA.

Email: harley@ncf.edu

The authors would like to thank…. Larry Boles, Lauren Puishys, Leah Maurer, Kara Watts, Mandi Schook, M. Andy Stamper, Kim Odell, Cathy Goonen, Lisa Green, Leslie Larsen-Plott, Seas interns…

Abstract (150-250 words)

Dolphins gain information through echolocation, a publicly accessible sensory system in which dolphins produce clicks and process returning echoes, thereby both investigating and contributing to auditory scenes. How their prior knowledge of these scenes contributes to their echoic information-seeking is unclear. Here, we investigate their top-down cognitive processes in an echoic matching-to-sample task in which targets and auditory scenes vary in their decipherability and shift from being completely unfamiliar to familiar. A blind-folded adult male dolphin investigated a target sample positioned in front of a hydrophone to allow recording of clicks, a measure of information-seeking; the dolphin received fish for choosing an object identical to the sample from 3 alternatives. We presented 20 3-object sets, unfamiliar in the first of 5 18-trial sessions with each set. Performance accuracy and click counts varied widely across sets. Click counts of the 4 lowest-performance-accuracy/low-discriminability sets (X=43%) and the 4 highest-performance-accuracy/high-discriminability sets (X=91%) were similar at the start of the first session and then decreased for both kinds of scenes, although the decrease was substantially greater for the low-discriminability sets. In 4 challenging-but-doable sets, number of clicks remained relatively steady across the 5 sessions. Reduced echoic effort with low discriminability sets was not due to overall motivation: the differential relationship between click number and object set discriminability was maintained when difficult and easy trials were interleaved and when objects from originally difficult scenes were grouped with more discriminable objects. These data suggest dolphins calibrate their echoic information-seeking effort based on their knowledge and expectations of auditory scenes.

Keywords: marine mammals, dolphins, echolocation, auditory scenes, top-down processing (4-6)

Information-seeking across Auditory Scenes by an Echolocating Dolphin

The world can be a “blooming, buzzing confusion” offering myriad sources of detectable information to the sensory systems of an organism trying to negotiate its world (James, 1890). Determining how organisms manage the challenge of organizing this input has been a central focus in psychology since its inception – and before. In his *Principles of Psychology*, James began his chapter on discrimination by citing the 17th-century philosopher Locke: “It is not enough to have a confused perception of something in general: unless the mind had a distinct perception of different objects and their qualities, it would be capable of very little knowledge”. Indeed, object perception and recognition are central to creating an animal’s representation of its world, its Umwelt. Under the sea, parsing the blooming, buzzing confusion into objects might be especially difficult due, in part, to the efficacy of the transmission of sound in water (sound travels more than four times faster in water than in air) and the water’s frequent murkiness, which limits vision. However, bottlenose dolphins (*Tursiops truncatus*, hereafter referred to as dolphins) evolved to take advantage of sound’s underwater speed through the development of echolocation, a system by which they send short, focused, high intensity, broadband clicks into the water and interpret the returning echoes to detect, discriminate, and identify objects, the surface, and other aspects of the auditory scene (e.g., Au, 1993; Nachtigall & Moore, 1988; Harley, Putman, & Roitblat, 2003; Pack & Herman, 1995; Xitco & Roitblat, 1996; Houser, Martin, Bauer, Phillips, Herrin, Cross, Vidal, & Moore, 2005). What is still unclear is how dolphins manage their echoic object recognition. Here, we try to gain a better understanding of the dolphin’s world by studying the effects of top-down processing (knowledge, memory, expectations, selective attention) on echoic information-seeking of unfamiliar-to-familiar targets ultimately presented in auditory scenes that vary in their difficulty.

Most studies of dolphin echolocation have focused on bottom-up processes associated with echolocation, and although a great deal of important psychophysical research has uncovered some of the exquisite discrimination abilities dolphins have when it comes to detecting acoustic differences in frequency (Thompson & Herman, 1975), amplitude (Evans, 1973; Au, 1993), and time (Au, Moore, & Pawlowski, 1988; Moore, Hall, Friedl, & Nachtigall, 1984), these animals must surely also rely heavily on top-down processes to organize and interpret the echoes they receive which are an inconsistent information source for multiple reasons. Firstly, sound is affected by many physical marine attributes – air bubbles, temperature, salinity, density, surface characteristics, and more – likely requiring the dolphin to use its knowledge of these forms of interference as well as its knowledge of familiar objects to manage object recognition. Secondly, objects are complex as the effect that echoes from different aspects of the same object can vary more than those from different objects (DeLong, Au, Lemonds, Harley, & Roitblat, 2006), and similar object features can produce different sounds depending on many factors including physical conditions and angle of reflection (Helweg et al., 1996). Thirdly, the larger auditory scene is formidable, with echoes returning from multiple objects including fish, rocks, conspecifics, and air (e.g., the water's surface), as well as having interference from sounds issued by animals of all kinds against the backdrop of the many varieties of the ocean's roar. With that in mind, one wonders if the ability to hear from below 1 kHz to above 150 kHz (Johnson, 1967; Houser & Finneran, 2006) is a blessing or a curse!

For object recognition, the dolphin's broad frequency hearing range does in fact help, as do its other perceptual strengths. Interpretations of dolphins' object/echo discrimination performance accuracy and confusions coupled with analyses of objects' echoes and perceptual models suggest that dolphins may use multiple echoic attributes to recognize objects including time separation pitch (a perceived frequency created by the return of correlated sound pulses received closely in time), target strength (the intensity of echo returns from target objects), the pattern of changes of echoes from multiple object orientations, the distribution of energy across frequencies, peak frequency, center frequency, duration, and integration of echoes, among others, and that they likely use a combination of attributes rather than single attributes (e.g., Au & Martin, 1989; DeLong, Au, Lemonds, Harley, & Roitblat, 2006; Helweg, Au, Roiblat & Nachtigall, 1996; Helweg, Roitblat, Nachtigall, & Hautus, 1996; Branstetter, Van Alstyne, Strahan, Tormey, Wu, Breitenstein, Houser, Finneran, & Xitco, 2020; Altes, Dankeiwicz, & Moore, 2003). But sounds alone do not a perceptual object make. The dolphin's perceptual and cognitive systems do that work: How?

Auditory Scene Analysis

A general framework to help us think about this problem is Bregman's (1990) auditory scene analysis, an approach focusing on how humans and other animals organize apparent cacophony, i.e., how they group sounds from the same source together into auditory streams and disambiguate/separate those streams from sounds related to other sources (Bregman, 1990, 2005, 2015). Bregman uses the Gestalt approach to visual processing as an analogy: In the same way that people use the principle of similarity to group dots of the same color into columns when they appear on a page full of lines of equidistant dots in which every other dot in each row alternates from yellow to blue to yellow to blue, etc., people also group a series of alternating high and low tones into two auditory streams: a high one and a low one. An everyday example of the organization of auditory objects can be illustrated in the well-known example of the phonemic restoration effect in which listeners fill in speech sounds that are missing due to the inclusion of coughs or other noises that replaced the expected speech sound. This phenomenon occurs through a combination of bottom-up (accessing the speech sounds themselves) and top-down (filling in information based on knowledge, expectations, and goals) processes applied when the sounds that listeners hear are ambiguous and they are trying to organize an auditory scene (e.g., Shinn-Cunningham & Wang, 2008). Top-down processes not only fill in missing information, but also guide information-seeking behavior. For example, listeners can identify auditory objects based on timing or frequency attributes, a bottom-up process, but providing instructions to them can sway them to attend to one attribute versus another, a top-down process, and these changes in attention affect the auditory object they perceive (Bregman, 1990; Shinn-Cunningham, Lee, & Oxenham, 2007; Shinn-Cunningham, 2008; Miller & Bee, 2012).

The use of top-down processes typically makes navigating the world faster and easier. Although learning requires time and resources, having knowledge reduces the need for resources later and makes problem solving (e.g., the problem of identifying words, objects, and patterns) easier in a complex and noisy world. For example, memory of previous familiar problems and patterns is central to problem solving for experts like chess masters and makes them significantly faster and more capable at coming to successful solutions than novices (see Bilalic, McLeod, & Gobet, 2009, for an overview). This reduction in the use of cognitive resources after learning is also reflected in neural processing in multiple ways including via long-term potentiation, a strengthening between neural synapses occurring after repeated activation of that neural channel (see Hayashi, 2021, for an overview). Trying to avoid a predator or capture prey? Faster identification is better! For an echolocating dolphin moving through the world, decoding auditory scenes in part through object recognition should require fewer resources for familiar objects and scenes, thereby making life more manageable.

The framework of auditory scene analysis applied to other animals, including macaques, starlings, treefrogs, ferrets, budgerigars, bats, and dolphins indicates that they also organize their acoustic environment into informative auditory streams (Bee, 2015; Itatani & Klump, 2009, 2020; Ma, Micheyl, Yin, Oxenham, & Shamma, 2010; Neilans & Dent, 2015; Fishman, Reser, Arezzo, & Steinschneider, 2001; Fishman, Micheyl, & Steinschneider, 2012; Hulse, MacDougall-Schackleton, & Wisniewski, 1997; Moss & Surlykke, 2001; Branstetter & Finneran, 2008; Branstetter, Trickey, Bakhtiari, Black, Aihara, & Finneran, 2013; Finneran & Branstetter, 2013). Work with bats is of special interest here because they, like dolphins, are also echolocators and can directly contribute to the auditory scenes they are decoding through their biosonar systems, i.e., how they manage their echoic investigations in part determines the echoes they receive. Bat echolocation systems have evolved for the different needs and habitats experienced across the hundreds of different echolocating bat species on the planet, and individual bats can adaptively control their outgoing echolocation signals and movements to gain better information about the world and likely reduce the complexity of an auditory scene (Moss & Surlykke, 2010). For example, some species shift the frequency of their calls when in groups, presumably to help them disambiguate their own echoic returns versus those of a conspecific, and other species appear to avoid simultaneously echolocating with nearby bats (as do rough-toothed dolphins, Gotz, Verfuss, & Schnitzler, 2006) (Chiu et al., 2008, 2009; Jarvis et al., 2010).

Moss and her colleagues (e.g., 2010, 2016, 2018) make compelling arguments that bats' information-seeking behaviors, i.e., their active-sensing movements and behavioral responses to a variety of acoustic scenes, provide useful windows into bats' echolocation systems that enhance our understanding of the information bats are trying to control within the acoustic streams they are receiving. For example, when capturing prey in open rooms, a fairly easy auditory scene, versus near vegetation, a more difficult auditory scene, big brown bats change a host of their behaviors: When the prey are near vegetation, the bats take longer to try to intercept the prey target, spend more time "strobing" (producing packets of pulses with a stable pulse interval), increase the pulse intervals in the strobes potentially to increase processing time for this difficult task, change their flight paths to avoid the backward masking of echoes returning from the vegetation and overwhelming the target's echo, shorten the length of the terminal buzz (the final burst of pulses) produced just before capture, and are significantly less likely to try to capture the target in the first place than they do in an open room (Moss et al., 2006). Similarly, barbastelle bats shift their echolocation behaviors based on the difficulty of different tasks including increasing their call rates (i.e., effort) for more difficult tasks (Lewanzik & Goerlitz, 2021). Bats clearly adapt their echolocation behaviors literally on the fly to enhance perception and reach their goals, and analyzing those behaviors gives us insights into the information the bats need and how they work to get it across contexts.

Echoic Information-seeking by Dolphins: Strategies, Top-down Processing, & Effort

Although the general lack of availability of dolphins for research and the challenges of working in a marine environment have led to many fewer studies of free-swimming dolphins specifically focused on their echoic investigations of objects and auditory scene compared to bats (Moore & Finneran, 2011; Moss, Chiu, & Moore, 2014), many experiments with stationary dolphins and a few free-swimming animals indicate that they can control their echoic investigations of objects, presumably to get better information, including by producing more intense clicks with higher peak frequencies in noise, emitting more clicks, and getting closer to objects when they are free-swimming (e.g., Au et al., 1974; Au et al., 1982; Houser, Martin, Bauer, Phillips, Herrin, Cross, Vidall, & Moore, 2005; Roitblat, Penner, & Nachtigall, 1990). We also know that free-swimming echolocating harbor porpoises (*Phocoena phocoena*) change their investigatory behaviors when auditory scenes change: When a target object and an alternative were closer together, potentially making an auditory scene more difficult, the porpoises increased the number of their echoic scans, clicked faster, began their buzzes farther away and increased their duration, and made their decisions when they were closer to the targets (Malinka, Rojano-Donate, & Madsen, 2021).

Some data shed light on the top-down processing mechanisms bottlenose dolphins use for echolocation tasks. For one, dolphins have remarkable echoic attention capacities. Because dolphins are unihemispheric sleepers, they can echoically monitor their environment for echoic targets for at least 15 continuous days, i.e., 360 hours in a row, with high performance accuracy (>95% with an average of 78.4 trials/day) (Branstetter, Finneran, Fletcher, Weisman, & Ridgway, 2012). For another, expectations can affect click production; dolphins use their expectations of a target's distance to space their clicks (Au, Floyd, Penner, & Murchison, 1974; Au, Penner, & Kadane, 1982). For example, in a task in which an echolocating dolphin had to report the presence or absence of an object at 5 different distances, the dolphin performed significantly better when the object appeared at the same distance throughout a session than when it appeared at different distances, suggesting that focusing on a single distance helped (Penner, 1988). In addition, the dolphin's inter-click interval was appropriate to a specific single distance (based on two-travel time of the out-going click and the returning echo) on both presence and absence trials when distance stayed the same throughout a session, indicating that the dolphin's knowledge dictated the nature of his click trains. Thirdly, object familiarity makes a difference. Dolphins in matching tasks often get better at matching as objects become more familiar (e.g., Herman, Pack, & Hoffmann-Kuhnt, 1998; Xitco & Roitblat, 1996).

The object-familiarity edge suggests that dolphins remember an object's echoes. This may mean that they need less information to recognize an object thereby allowing them to recognize objects more capably in contexts in which echoes are degraded or otherwise less accessible. They may also learn to investigate objects more efficiently in order to discover the telling characteristic of that object compared to alternatives. An unusual study on eavesdropping in dolphins adds support to these possibilities. In this study, a non-echolocating dolphin (the "listener") listened in on the echoes returning to its echolocating partner (the "inspector"). Both dolphins were originally trained to engage in an active echoic matching task and improved with object familiarity, indicating they remembered the objects. The listener tended to be better at these tasks; the inspector had more biases. When just eavesdropping, the listener's performance accuracy was above chance levels and numerous analysis indicated that his choices were not contingent on the inspector's choices; that is, echoic returns of the sample target object were the basis of his choices. However, the listener was affected by the inspector's investigations: The listener was more likely to be correct when the inspector was correct, the inspector's biases were reflected in the listener's choices (i.e., the inspector had a bias towards wooden and styrofoam objects, and when eavesdropping – and only when eavesdropping – so did the listener), and the inspector – and therefore the listener – was much more likely to be correct when he was inspecting familiar objects versus objects that were only familiar to the listener.

Effort also varies across echoic investigations. In one study of a pair of free-swimming dolphins' echoic behaviors when engaging in a presence/absence target detection task in a cluttered environment, one "methodological and thorough" dolphin (whose high-frequency hearing was likely significantly worse than the second dolphin's) averaged over 300 clicks and 25 seconds in a target-present search and over 500 clicks and 38 seconds in a target-absent search, whereas the second minimalist dolphin averaged around 31 clicks and 6.5 seconds for target-present and 109 clicks and 18 seconds for target-absent searches (Houser et al., 2005). Both dolphins increased the strength of echo returns at target acquisition, the first dolphin by getting closer to the target, the second dolphin by increasing the intensity of his clicks. In an early study of effort, two stationary echolocating dolphins in a presence/absence target detection task were more similar to each other than in the free-swimming example, but also interesting because in this task the noise level was increased across trials until the two animals eventually gave up (Au, Penner, & Kadane, 1982). The mean number of clicks per trial began between 20-30 clicks/trial and, as the noise level was increased, number of clicks steadily increased 2-3 fold until a noise threshold was reached at which point number of clicks went steadily down; in fact, one dolphin avoided clicking altogether for some of the high-noise trials. Apparently, when the problem became insoluble, the dolphins stopped trying to solve it.

In an analysis of problem solving involving decision-making and object recognition in a stationary echolocating dolphin performing a matching-to-sample task, Roitblat, Nachtigall, and Penner (1991) (like others) framed the matching paradigm as two discrimination problems, a successive discrimination problem to identify the sample target object and a second simultaneous discrimination problem to choose the matching choice from the group of alternative objects, a simple auditory scene. Both problems require sequential-decision sampling: Identifying the sample requires repeated effortful investigations that build up information over time; determining the object that matches the sample from an array of objects is less effortful because it only requires gaining enough information to determine similarity to the sample. The highly skilled dolphin Rake, almost error-free with his very familiar four objects, varied his effort in investigations. As expected he invested more effort in identifying the sample target objects (averaging 37.2 echoes) than in his first scan to the alternative objects (averaging 17.3 clicks). He also traded effort for accuracy, working harder to identify the weak-return most difficult objects (the most missed), averaging more clicks to the matching stimuli (versus the non-matching stimuli) in his first scan of the alternative array, and extending his efforts to match alternatives on the right which took more scans since he investigated the alternatives stereotypically from left to right. Roitblat et al. also used their data to create a decision-making model enlisting signal detection theory and Bayesian decision rules that indicated the dolphin was integrating information from successive investigations to inform his echoic information-seeking behaviors. In this model, dolphins gain more knowledge with more clicks, and recognition of the sample target requires the most effort because it requires full recognition rather than merely evaluating for similarities between the target and the choice.

The Current Study

Auditory scene analysis embraces the whole blooming, buzzing confusion of an animal’s acoustic world, and because echolocators orchestrate aspects of the composition of that scene through their clicks/pulses and movements to decode the returning echoes, their information-seeking behaviors also give researchers information about what the animals need and use to create auditory objects. When processing a scene, animals may also assess the scene for its decipherability and calibrate their efforts based upon the possibilities, as the dolphins in the object detection task outlined above did when noise overwhelmed the object’s returning echoes. Using the famous everyday-life example of tracking a speaker’s words at a cocktail party, the acoustic scene may allow the listener to hear easily, lead the listener to put in extra effort to manage information exchange, or require the listener to expend so much effort she gives up or invites the speaker to an easier venue (a different scene), all depending upon multiple characteristics related to the scene itself, the sound source, and the listener.

When dolphins echolocate, they vary their investigations in order to get better information, but specifics of how and why are still emerging from the noise. Target discriminability and familiarity appear to have an influence on dolphins' echoic investigations of objects, but to date there have been no studies that systematically examine the influence of these variables in an echoic object recognition task. Our general goal here is to learn more about how a dolphin's top-down cognitive processes affect his responses to different auditory scenes as they become more familiar. We begin simply, by measuring the effort a free-swimming dolphin invests as he echolocates target samples that appear in fairly straight-forward acoustic scenes (the three alternatives that appear in choice arrays in a matching-to-sample paradigm) in which the discriminability of the objects themselves makes recognition harder or easier to see if the dolphin takes these variations in decipherability of the scenes into account. This approach allows us to explicitly look at the top-down processes of attention, memory, knowledge, and expectation as the targets and scenes methodically shift from being completely unfamiliar to familiar. Our specific question for this study is: Does a dolphin's echoic investigatory practices, in this case the number of clicks he produces to target sample objects, change as the dolphin learns about auditory scenes that range in their decipherability? Given the dolphin's prowess at echoic object recognition and the variability of echoes across auditory scenes, we predict that the dolphin's investigations of objects will change as it learns more about objects within specific contexts, i.e., it will bring top-down processes to bear to drive its information-seeking behavior.

Methods

Subject

The subject was an adult male bottlenose dolphin (*Tursiops truncatus*), Calvin, who had previous visual and echoic matching-to-sample experience (Harley, Fellner, & Stamper, 2010). He was born in 1994 at a facility in the Florida Keys and moved to his current facility in 2003. Audiograms in 2013 and 2019 conducted via auditory evoked potential confirmed that he had good hearing across the normal range. He lived with 3 other adult males in one quarter+ of a 5.8 million-gallon, mixed-species exhibit and two ancillary pools at The Seas, Epcot®, Walt Disney World® Resort, Lake Buena Vista, FL, USA. The current study was conducted in one of the ancillary pools (B Pool) which measured 8.2m long by 7m wide by 2.1m deep. See Figure 1 for the study location, B Pool. Calvin consumed a diet of herring, capelin, and squid that was customized by nutritionists, veterinarians, and trainers on the Animal Health and Animal Care teams. These teams were responsible for all care and management decisions which were independent of Calvin's participation or accuracy in research sessions. Disney's Animal Care and Welfare committee reviewed and approved of the project (IR1005 and IR1809), and the dolphin was cared for in accordance with the U.S. Animal Welfare Act, the Marine Mammal Act (1972), and the Association of Zoos and Aquariums (2014) accreditation guidelines at all times. The Seas was authorized to house the animals by permit # 58-C-0076 issued by the U.S. Department of Agriculture Animal and Plant Health Inspection Service.

Because dolphins are difficult to access, and tasks like this one take extensive periods to train, one dolphin participated in the study. Single-subject designs do not allow for group comparisons, but our question focused on the capacity of any dolphin to use top-down processes to calibrate its information-seeking behaviors based on auditory scenes. Therefore, working with a single subject allowed us answer our question and benefit from a fine-tuned analysis of the dolphin's responses to the independent variables of object familiarity and scene decipherability. In some circumstances, the individual level, using within-subject repeated sampling, is more effective than large numbers of subjects (Smith & Little, 2018). We designed our study to take advantage of this strength, due to access limitations and the nature of our question.

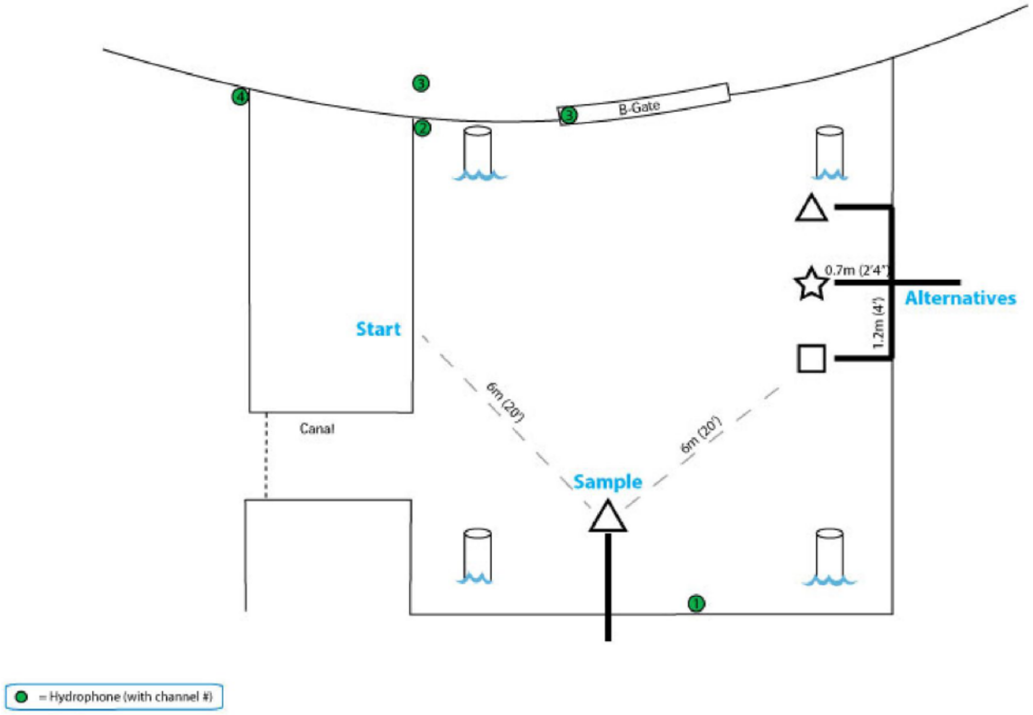


Fig. 1 The experimental set-up. The echolocating dolphin begins with the trainer at "start", swims to the sample object (hydrophone #1 positioned behind object), and proceeds to the alternative array to make a choice

Materials

Stimuli

The main focus of the study was to examine object familiarity across different levels of discriminability, and so many objects served as stimuli. In the beginning of the study, all stimuli were unfamiliar to the subject and were only used in test sessions, as described throughout the manuscript. (The dolphin learned the task in a previous study with other objects.) For the bulk of the study, we used 20 3-object sets (Sets A-T presented chronologically in alphabetical order), although there were a few other objects introduced in the last condition of the study. Most objects were made of PVC, but some were hardware “junk” objects made of a variety of materials. Across the study, the PVC objects ranged in size from 2.7” H by 4.3” (the smallest) to 30.8” H by 18.1” W (the largest). The hardware objects ranged in size from 2.6” H by 1.9” (the smallest) to 14.6” H by 12.8” W (the largest). The smallest hardware object was the Stapler Remover, and the largest hardware object was the Letter Tray. Figure 2 shows examples of object stimulus sets from the original 20 sets. Figure 3 shows examples of the objects used in the three trial types in the final condition of the study, “Scene Shifts”, in which familiar objects were inserted into new scenes with unfamiliar objects. Figure 4 shows examples of the unfamiliar easy “junk” hardware objects used in the “Scene Shifts” condition.

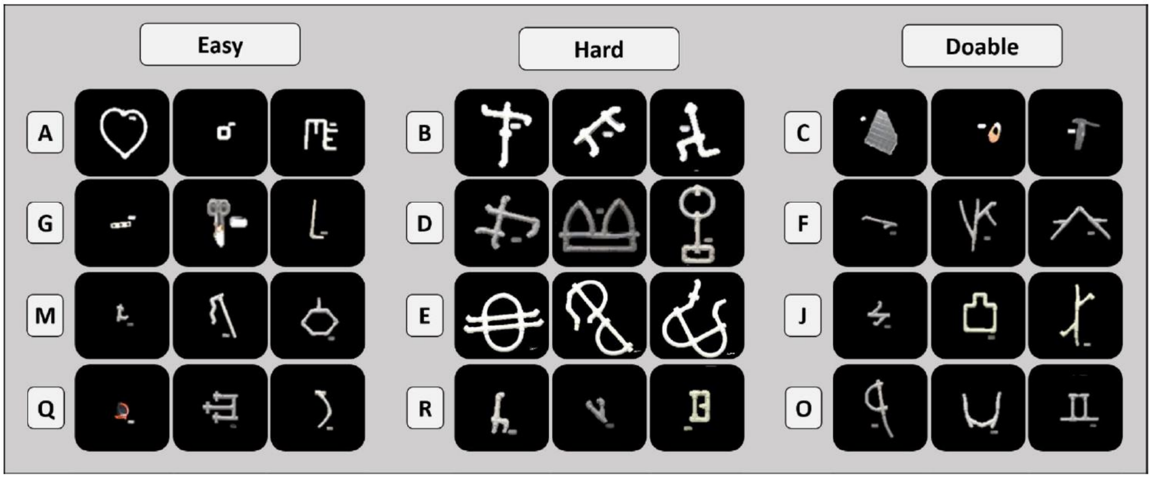


Fig. 2 Examples of stimulus sets: easy (first block, sets by row), hard (second block, sets by row), and challenging-but-doable (third block, sets by row)

The sample and alternative objects all hung off PVC poles via monofilament line covered in small clear soft-plastic tubes. All objects were suspended 0.7 m from the walls of the pool, with 1.2 m between each alternative. Objects were suspended such that the center of each object was 40.6 cm under the surface of the water.

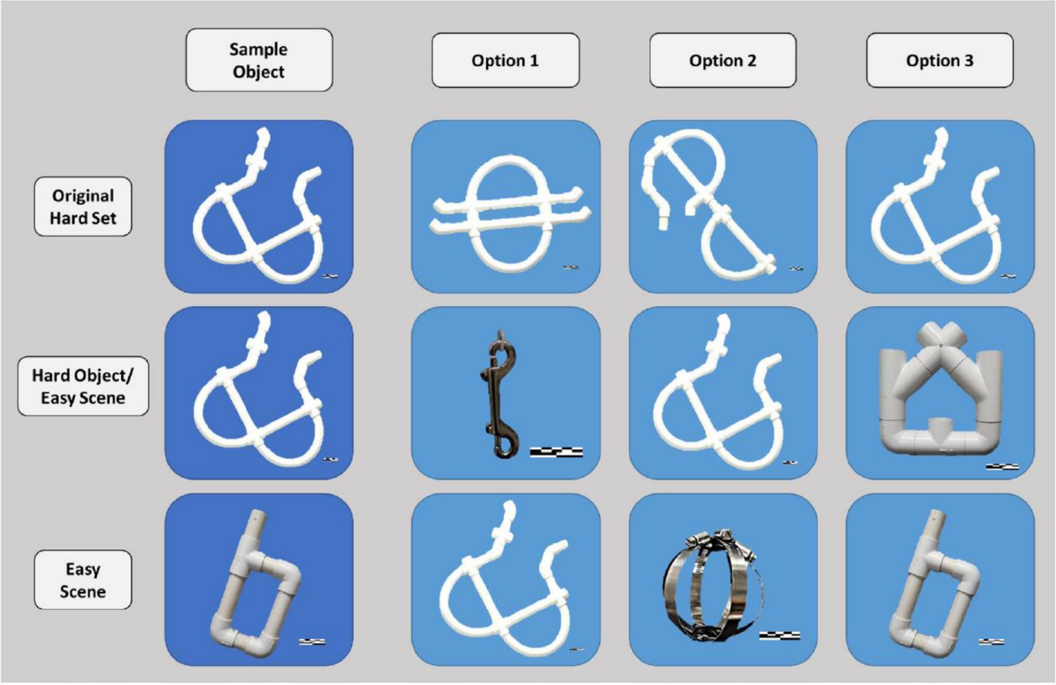


Fig. 3 Examples of the three trial types in the Scene Shifts condition featuring an original hard set (top row), a hard-set object presented within an easy scene (middle row), and an easy scene (bottom row)

Recording and Acoustical Analysis Devices

We made acoustic recordings during the sessions with High Tech, Inc. HTI-96MIN hydrophones with a flat frequency response of 2 Hz to 30 kHz (although the actual recording range was 0 Hz to 50 kHz), and clicks were recorded at a sampling rate of 100,000 Hz per second. The clicks were recorded onto a Lenovo T410 laptop computer using Avisoft-RECORDER USG version 4.2.8 (http://www.avisoft.com). During all sessions, a hydrophone was mounted directly behind the sample object to record Calvin’s clicks directed toward the target sample object. Aside from the sample hydrophone, there were two other hydrophones in the pool. The audio recordings using the three channels were analyzed with Avisoft-SASLab Pro Sound Analysis and Synthesis Laboratory version 5.2.01. See Figure 1 to see locations of the hydrophones.

Fig. 4 Examples of unfamiliar easy "junk" hardware objects used in the Scene Shifts condition

Video was recorded using a PC Osprey 4-channel video card with H.264 Webcam software that simultaneously recorded above the water in all the dolphin areas. The camera over B Pool, the study site, was mounted such that it could capture the entire pool.

Procedure

At the start of the study, Calvin already performed capably in a 3-alternative matching-to-sample task. He wore soft, latex eyecups during trials to preclude visual cues. He could pop the eyecups off at will, but he was trained to wear them. If an eyecup came off during a trial, stimuli were immediately pulled from the pool.

Before the start of a trial, the sample was positioned below the surface of the water; it was introduced along with a foil object to disguise any splash cues associated with putting the sample into the water, and the foil was then pulled out of the water leaving only the sample. The alternative choice objects were also positioned before the start of a trial. See Figure 1 for the trial set-up.

A trial began when the dolphin positioned himself in front of his trainer, who then signaled him tactilely to swim 6.1 m to the target sample object to his left. The dolphin could swim at his own pace. After investigating the sample object ad libitum, he swam to the choice array 6.1 m to his left where three alternative objects, one identical to the sample, were positioned. After inspecting the alternatives, Calvin positioned himself in front of his choice object and chirped. A research assistant who did not know the identity of the sample identified the object Calvin had chosen. If Calvin's choice matched the sample, the trainer blew a secondary reinforcer "bridge" whistle and gave Calvin 2-3 capelin. If his choice was incorrect, the trainer interacted briefly with Calvin, and we moved to the next trial. Occasionally, the trainer interacted informally with Calvin between trials. Intertrial intervals were a minimum of 30 seconds.

We recorded clicks to the sample by mounting the hydrophone directly behind the sample object. A researcher recording the session inserted comments into Avisoft SasLab at the moment the dolphin was: 1) released to approach the sample, 2) turned away from the sample toward the alternatives, and 3) chirped in front of one of the alternatives to indicate his choice. All clicks between the time the dolphin began approaching the sample and began turning away from the sample were counted in the click counts reported as *information-seeking clicks,* i.e., the clicks the dolphin emitted during the sample investigation period. The researcher's comments on start and end of Calvin's sample investigations were confirmed by a second researcher who synchronized the video and audio using a visible and audible synch tap and then used Avisoft SASLab's "pulse train analysis" to count the number of clicks directed at the sample hydrophone. If Calvin's rostrum was pointed away from the sample object and the clicks were louder on channels 2 or 3 than on channel 1 (the sample hydrophone), then those clicks were not counted in Calvin's investigation of the sample. Calvin's clicks toward the sample usually either stopped or paused before he turned to the alternatives. We also hand-counted the number of clicks in 3% of the trials in a selection of 8 sets (the most and least discriminable sets) using a semi-randomized procedure in which we counted trials that included each sample within a set, always in different sessions with that set. We determined that no clicks were missed by the automated system, including clicks in terminal buzzes (click sequences with short inter-click intervals occurring at the end of the investigations of the target sample objects). However, we did discover that there were some falsely-detected clicks (mostly echoes at the beginning or end of a click train when the inter-click intervals were long) that inflated our numbers by about 7%. However, this error rate occurred in clicks across targets and auditory scene types, thereby having little effect on relative comparisons. Because we were interested in relative numbers of clicks and we always used the same procedure to count clicks, any sample-hydrophone-directed clicks Calvin may have produced to navigate the pool as he swam to the sample were anticipated to be relatively consistent or consistently variable across the more than 1800 trials we recorded.

Across the experiment, the dolphin was tested with 20 object sets, each of which included 3 different, unfamiliar objects. Each set was presented for 5 18-trial sessions. Having 18 trials allowed us to present the sample an equal number of times (6) and the presentation of its match and non-matching alternatives an equal number of times in each position in each session. Having

5 sessions with each set allowed us to explore any changes in the dolphin's information-seeking behaviors as the objects changed from being unfamiliar to familiar across the 5 sessions. Not only did object sets vary on familiarity over time, they also varied in their discriminability. We let the dolphin indicate object discriminability by using the dolphin's performance accuracy at identifying the object that matched the sample within the 3-alternative choice array to determine this variable. In our tests, chance performance accuracy (33% based on 3 alternatives in the choice array) would be interpreted to mean the objects in that set were indiscriminable to the dolphin, whereas high performance accuracy, e.g., 94%, would suggest the objects in that set were highly discriminable to the dolphin.

Control "interleaved" sessions to evaluate the effect of general motivation versus auditory scene on performance accuracy and effort

We hypothesized that changes in performance accuracy (discriminability) and number of clicks (attentive effort) might vary across the auditory scenes/object sets based on (1) the dolphin's experience with those specific scenes/sets OR (2) the dolphin's general motivational state across the time period during which the 20 scenes/object sets were tested. In order to distinguish between these two explanations, we ran 15 total control sessions in which we interleaved trials from 3 difficult object sets (mean performance accuracy < 45%) and 3 easy sets (mean performance accuracy > 88%) within 5 36-trial sessions. That is, we organized an easy object set into 18 trials as previously described for test sessions, and we organized a difficult object set in the same way. Then, we interleaved those trials to create a single 36-trial session in which trials from both sets were randomly ordered resulting in the interleaving of hard and easy trials. We created 3 interleaved sets for 5 36-trial sessions each. We reasoned that if the dolphin's general motivational state was responsible for the differences across sets, then presenting the sets within a single session should even out the differences; that is, performance accuracy and number of clicks should be similar across all of the trials. On the other hand, if the dolphin's performance accuracy and information-seeking (number of clicks) were due to his learning about the object sets themselves, then his performance accuracy and number of clicks in these interleaved sets should be similar to his original performance accuracy and number of clicks when the sets were presented the first time within the series.

Scene Shift sessions to evaluate the effects of a change on auditory scenes

Discriminability is based on auditory scene. That is, objects may be difficult to recognize in some contexts and easy in others. In order to make a small in-road into understanding the nature of Calvin's memory of objects related to scene, we conducted 9 total sessions that included 3 trial types based on the objects that were presented in the alternative arrays. Trial Type 1 included samples and alternative arrays comprised of objects that continued to be presented in sets from the original 20 sets but were sets on which Calvin had poor performance accuracy, i.e., "hard" sets. Trial Type 2 included samples from the "hard" sets and alternative arrays in which those samples were presented with new unfamiliar objects anticipated to be easy for Calvin to discriminate. Trial Type 3 included samples and alternatives created with new PVC and "junk" hardware objects that were anticipated to be easy for Calvin to discriminate. Each trial type occurred 6 times in each session leading to 18 total trials. Three "hard" sets were tested across 3 18-trial sessions along with unfamiliar "easy" objects unique to each set. Therefore, there were only 18 trials total of each trial type for a given set of objects based on the three sessions with each combination set. See Figure 3 for examples of these trial types.

Results and Analyses

Variability of performance accuracy and number of clicks across all object sets

As intended, difficulty of interpreting the auditory scene, defined empirically using the dolphin's mean performance accuracy with each set across 5 18-trial sessions with that set, varied widely across the 20 object sets (A-T, alphabetical listing reflects chronological order of presentation). Chance performance accuracy was 33% because there were 3 alternatives in each choice array. Mean performance accuracy on the 20 sets across the 5 sessions that each set was initially presented to the dolphin ranged from 33.0% to 93.33% with an overall mean of 66.83% (Standard Deviation = 18.45%) for all sets. Performance accuracy on individual 18-trial sessions ranged from 22.22% to 100%.

Information-seeking effort, defined as the number of clicks directed to the sample object, also varied widely across the object sets. Mean number of clicks towards the sample objects presented within the 20 sets across the 5 sessions that each set was initially presented to the dolphin ranged from 229.67 clicks to 473.83 clicks with an overall mean of 353.91 clicks (Standard Deviation = 73.51 clicks) for all sets. The mean number of clicks on individual 18-trial sessions ranged from a low of 154.56 clicks to a high of 694.5 clicks. (One of the 100 sessions had no click counts due to equipment malfunction.)

Figures #5A and #5B provide a sense of the variability across object sets in terms of performance accuracy and mean number of clicks directed towards the sample objects, respectively, across all of the sessions for each of the 20 object sets (100 total sessions).

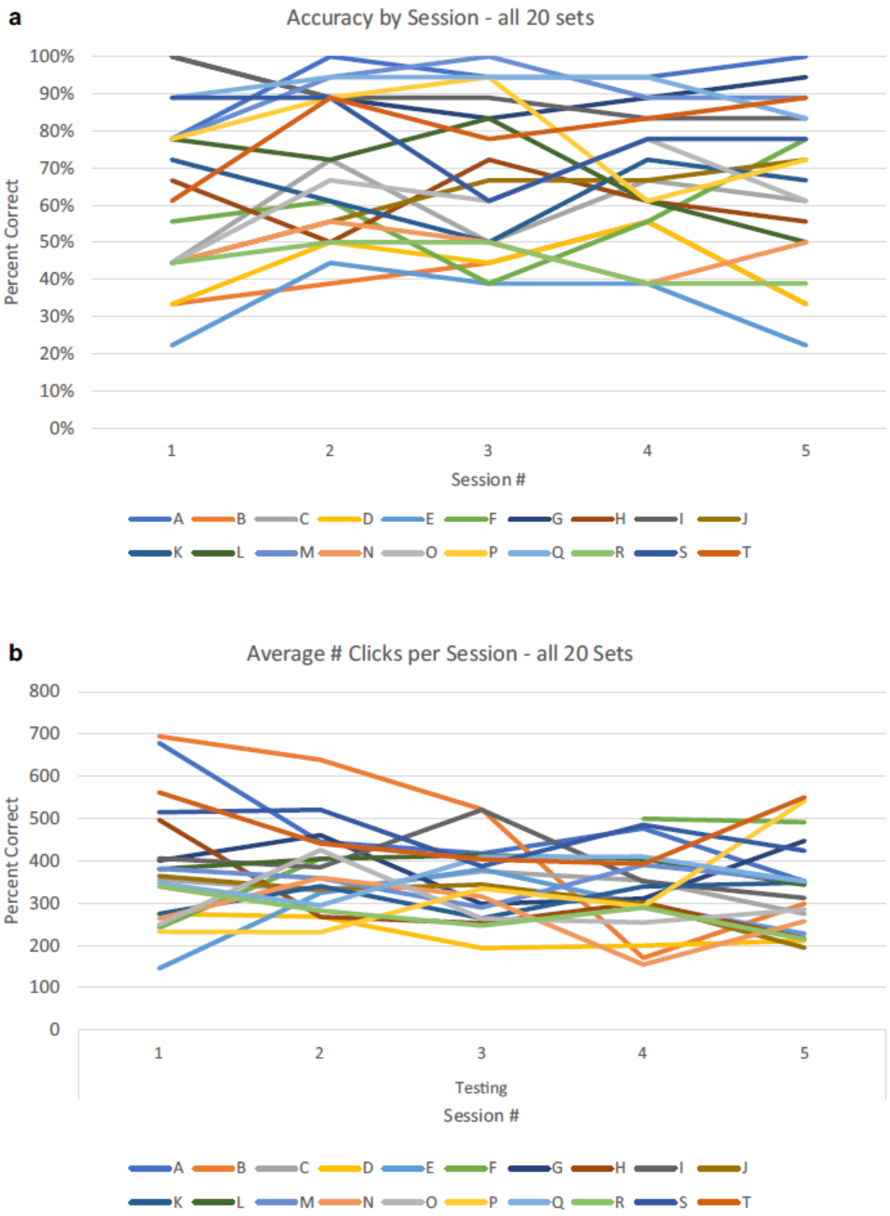
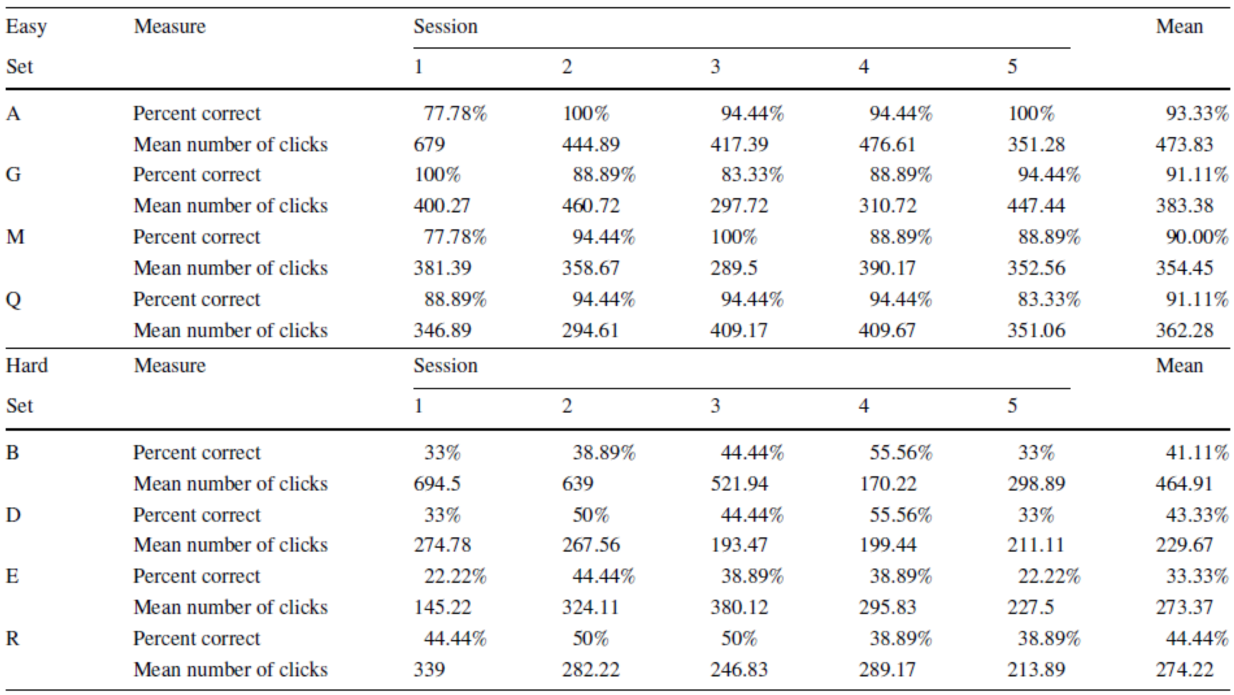


Fig. 5 **A** Accuracy by session for all 20 object sets, A-T (one line/set). **B** Average number of clicks/trial for each session for all 20 object sets, A-T (one line/set)

Information-seeking across object discriminability and familiarity

In order to learn more about the dolphin's information-seeking effort across the dimensions of object familiarity and discriminability, we focused several analyses on 8 object sets categorized as Easy and Hard as defined by the dolphin's mean performance accuracy in the matching task. The 4 sets with highest mean performance accuracy across the 5 sessions in which they were presented (range 90.0% - 93.33%) were designated as Easy sets. In these sets, 2 sets had 2 PVC object shapes and 1 junk hardware-store object, and 2 sets had 3 PVC object shapes. The 4 sets with lowest mean performance accuracy across the 5 sessions in which they were presented (range 33.33% - 44.44%) were designated as Hard sets and comprised of shapes built from PVC. See Figure 2 for examples of the stimulus objects.

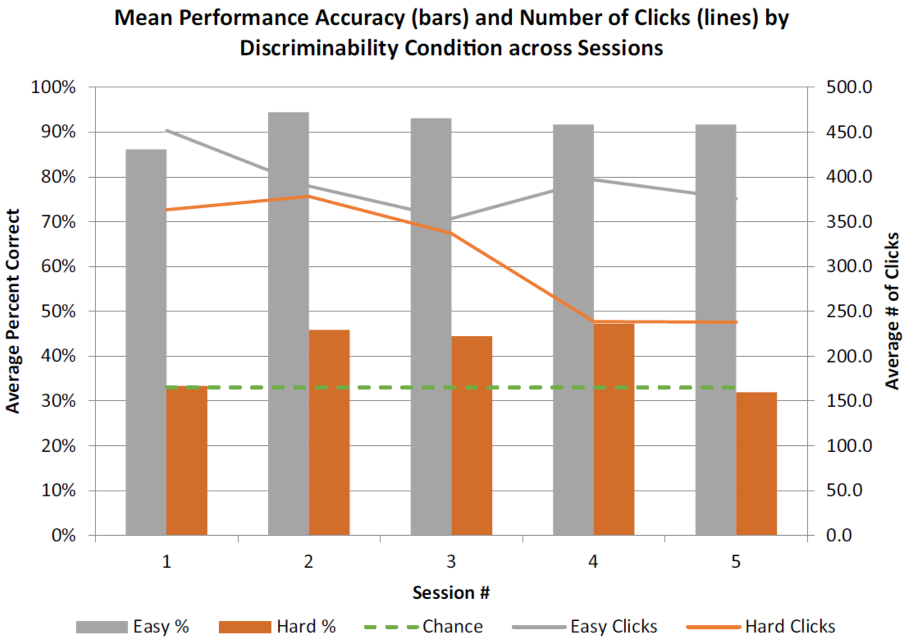
Table 1 Easy and hard sets: performance accuracy and mean number-of-clicks/trial by session



Easy Set	Measure	Session 1	2	3	4	5	Mean
A	Percent correct	77.78%	100%	94.44%	94.44%	100%	93.33%
	Mean number of clicks	679	444.89	417.39	476.61	351.28	473.83
G	Percent correct	100%	88.89%	83.33%	88.89%	94.44%	91.11%
	Mean number of clicks	400.27	460.72	297.72	310.72	447.44	383.38
M	Percent correct	77.78%	94.44%	100%	88.89%	88.89%	90.00%
	Mean number of clicks	381.39	358.67	289.5	390.17	352.56	354.45
Q	Percent correct	88.89%	94.44%	94.44%	94.44%	83.33%	91.11%
	Mean number of clicks	346.89	294.61	409.17	409.67	351.06	362.28
Hard Set	**Measure**	**Session 1**	**2**	**3**	**4**	**5**	**Mean**
B	Percent correct	33%	38.89%	44.44%	55.56%	33%	41.11%
	Mean number of clicks	694.5	639	521.94	170.22	298.89	464.91
D	Percent correct	33%	50%	44.44%	55.56%	33%	43.33%
	Mean number of clicks	274.78	267.56	193.47	199.44	211.11	229.67
E	Percent correct	22.22%	44.44%	38.89%	38.89%	22.22%	33.33%
	Mean number of clicks	145.22	324.11	380.12	295.83	227.5	273.37
R	Percent correct	44.44%	50%	50%	38.89%	38.89%	44.44%
	Mean number of clicks	339	282.22	246.83	289.17	213.89	274.22

Table 1 provides performance accuracy and mean number of clicks per trial in a session for the 8 object sets. As planned, the dolphin's performance accuracy was significantly better on the easy sets (4 sets of 5 sessions each: mean = 91.39%, SD = 6.85%) than hard sets (4 sets of 5 sessions each: mean = 40.56%, SD = 9.38%), paired $t(19) = 21.86$, $p < .00001$). The dolphin also investigated the sample object with significantly more clicks in easy sets (mean = 393.49 clicks, SD = 85.16) than hard sets (mean = 310.74 clicks, SD = 143.84), paired $t(19) = 2.99$, $p = 0.007$. (The Bonferroni correction to protect experiment-wise error at $p = 0.05$ with 7 tests is $p = 0.007$). See Figure 6 to see a graphic representation of these data.

Fig. 6 Mean performance accuracy (bars) and mean number of clicks/trial (lines) for sessions 1–5 for the 4 easy (first bar/pair; top line) and 4 hard sets (second bar/pair; bottom line)



All objects were initially unfamiliar to the dolphin at the beginning of the first session and thus became more familiar across the first and following sessions. Therefore, we compared the first and final sessions of the easy and hard object sets in terms of effort (number of clicks) reasoning that in the first sessions with each set the dolphin would be learning about these initially unfamiliar objects and their scenes, and that by the fifth sessions, the objects and their scenes would be familiar. Figures 7A and 7B show the mean number of clicks directed to the sample in each trial across the easy and hard object sets as well as performance accuracy on these trials in the first session with the object sets and the last session with the object sets, respectively. With the easy objects, the dolphin reduced the number of clicks he directed towards the sample from the first session with each set (4 18-trial sessions: mean = 451.89 clicks, SD = 178.45) to the fifth session with each set (4 18-trial sessions: mean = 375.58 clicks, SD = 120.05), paired $t(71) = 2.80$, $p = 0.007$. Similarly, though more pronounced with the hard objects, the dolphin reduced the number of clicks he directed towards the sample from the first session with each set (4 18-trial sessions: mean = 363.38 clicks, SD = 244.56) to the fifth session with each set (4 18-trial sessions: mean = 237.85 clicks, SD = 105.40), paired *t*(71) = 4.46, $p < 0.0003$. Overall, he dropped his effort to 83% of his original effort with easy sets and to 65% of his original effort with hard sets.

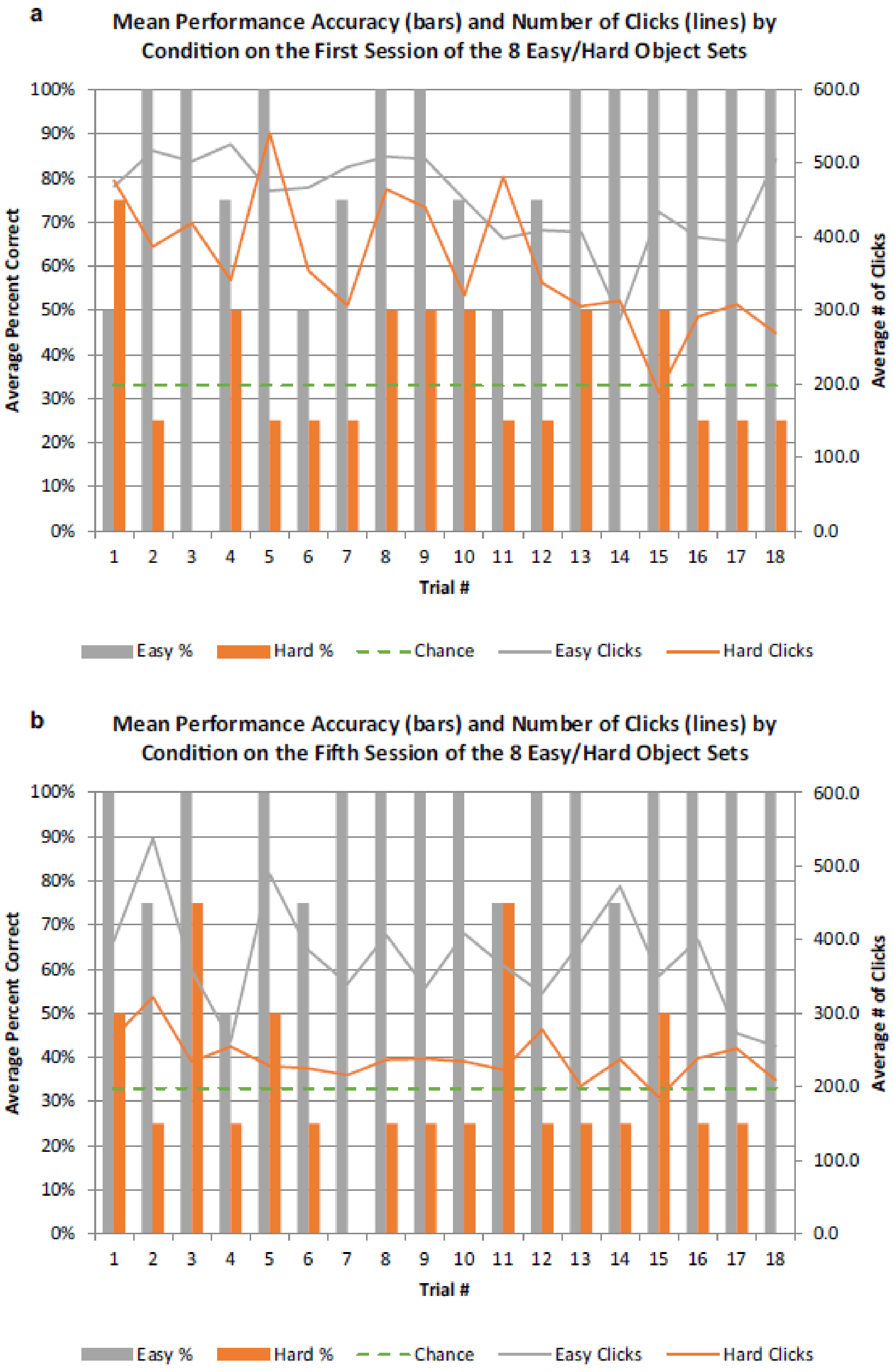


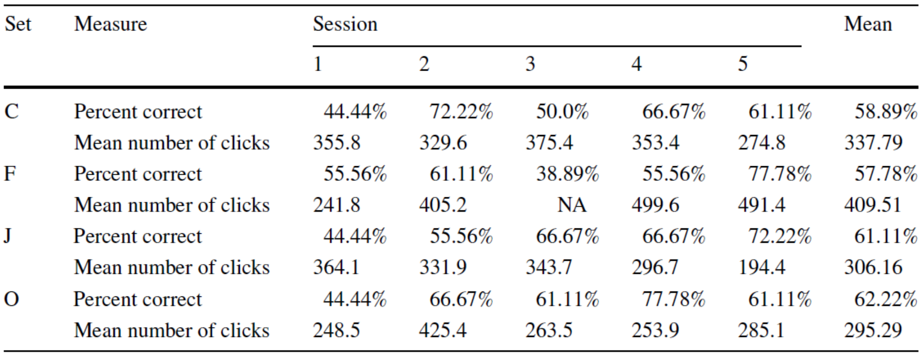
Fig. 7 **A** Mean performance accuracy (bars) and mean number of clicks/trial (lines) for the first sessions, trial by trial, of the 4 easy (first bars/pair; top line) and 4 hard sets (second bar/pair; bottom line). **B** Mean performance accuracy (bars) and mean number of clicks/trial (lines) for the fifth sessions, trial by trial, of the 4 easy (first bars/pair; top line) and 4 hard sets (second bar/pair; bottom line)

In order to get a more fine-tuned indication of how the dolphin's investigations of the objects changed as they became more familiar, we focused on the first trials and sessions with the objects. The dolphin had experienced each of the objects as a sample by the end of the first three trials of each first session. For easy sets, he produced 5949 clicks in those first three trials in the first session and 5176 clicks in the first three trials of the fifth session added across all 4 sessions with each easy set. For hard sets, he produced 5127 clicks in those first three trials in the first session and 3292 clicks in the first three trials of the fifth session. Hence, there was a 12.99% drop from first to last session in easy trials, and a 35.79% drop for hard trials; that is, the drop in effort for hard scenes was 3 times that of easy scenes for the first three trials of those sessions. Across all the first 8 18-trial sessions for easy and hard sets, the dolphin was incorrect on 10/72 easy trials and on 49/72 hard trials making his easy error rate 14.0% in the first easy sessions and his hard error rate 68.06% in the first hard sessions. Given these error rates, the discriminability of the objects within their auditory scenes was likely detected early in the first sessions.

Investigatory effort continued to shift across the five sessions in the Easy and Hard sets. As noted earlier, by the end of the first session, the dolphin's investigatory clicks were clearly reduced for hard sets ($\bar{X}$ per trial across the first session with each object set = 363.38 clicks) versus easy sets ($\bar{X}$ = 451.89 clicks), however, the number of clicks converged more closely in the second ($\bar{X}$ easy = 389.72 clicks; $\bar{X}$ hard = 378.22) and third ($\bar{X}$ easy = 353.44; $\bar{X}$ hard = 335.59) sessions and then diverged in the fourth ($\bar{X}$ easy = 396.79; $\bar{X}$ hard = 238.67) and fifth ($\bar{X}$ easy = 375.48; $\bar{X}$ hard = 237.85) sessions. Figure 6 illustrates the changes in investigatory effort and performance accuracy across the five sessions by easy and hard sets.

In contrast to Easy and Hard sets, investigatory effort did not significantly change for Challenging-but-doable objects sets across the 5 sessions. After discovering that number of clicks went down differentially on Easy and Hard sets across the 5 sessions with each set, we analyzed changes in effort for 4 Challenging-but-doable sets (C, F, J, O) as a way to tease apart the effects of familiarity and scene decipherability on effort. We chose the 4 sets in which the dolphin's performance accuracy was low in the first session ($\bar{X}$ = 47.22%, ranging from 44.44% to 55.56%) but had improved at least 15% by the fifth session ($\bar{X}$ = 68.06%, ranging from 61.11% to 77.78%). We discovered that mean number of clicks/trial with these sets, in contrast to the Easy and Hard sets, did not change significantly from the first ($\bar{X}$ = 302.56 clicks, SD = 121.53) to the fifth ($\bar{X}$ = 311.42 clicks, SD = 149.13) sessions, paired $t(71) = -0.3499$, $p = 0.727$. See Table 2 for performance accuracy and mean number of clicks/trial for each session for these rather variable sets.

Table 2 Challenging-but-doable sets: performance accuracy and mean number-of-clicks/trial by session



Set	Measure	Session					Mean
		1	2	3	4	5	
C	Percent correct	44.44%	72.22%	50.0%	66.67%	61.11%	58.89%
	Mean number of clicks	355.8	329.6	375.4	353.4	274.8	337.79
F	Percent correct	55.56%	61.11%	38.89%	55.56%	77.78%	57.78%
	Mean number of clicks	241.8	405.2	NA	499.6	491.4	409.51
J	Percent correct	44.44%	55.56%	66.67%	66.67%	72.22%	61.11%
	Mean number of clicks	364.1	331.9	343.7	296.7	194.4	306.16
O	Percent correct	44.44%	66.67%	61.11%	77.78%	61.11%	62.22%
	Mean number of clicks	248.5	425.4	263.5	253.9	285.1	295.29

Evaluating the effects of general motivation versus auditory scenes on performance accuracy and effort through "interleaved" sessions

We confirmed that the differences in performance accuracy across Easy and Hard sets were not due to variability in general motivation for echoic object recognition tasks via an analysis of the interleaved control sessions in which 2 18-trial easy and hard session trials occurred in random order within a single 36-trial session for 5 sessions per 6 combined object sets (Interleave Set 1 combined original Easy set G and Hard set D; Interleave Set 2: Easy I and Hard E; Interleave Set 3: Easy Q and Hard R). Performance accuracy (original $\bar{X}$ easy = 90.37 & interleaved $\bar{X}$ easy = 94.07%; original $\bar{X}$ hard = 40.37% & interleaved $\bar{X}$ hard = 43.33%) rose a little with these familiar objects but was not significantly different between the original and interleaved sets, paired $t(29) = -1.46$, $p = 0.155$. However, number of clicks (original $\bar{X}$ easy = 380.2 clicks & interleaved $\bar{X}$ easy = 404.2 clicks; original $\bar{X}$ hard = 259.1 clicks & interleaved $\bar{X}$ hard = 323.8 clicks) rose in the interleaved sets compared to the original sets, paired $t(29) = -2.33$, $p = 0.027$. (Note that the Bonferroni correction to protect experiment-wise error at $p = 0.05$ with 7 tests is $p = 0.007$). Table 3 presents mean performance accuracy and number of clicks/trial for the object sets used in this analysis when originally presented and later interleaved.

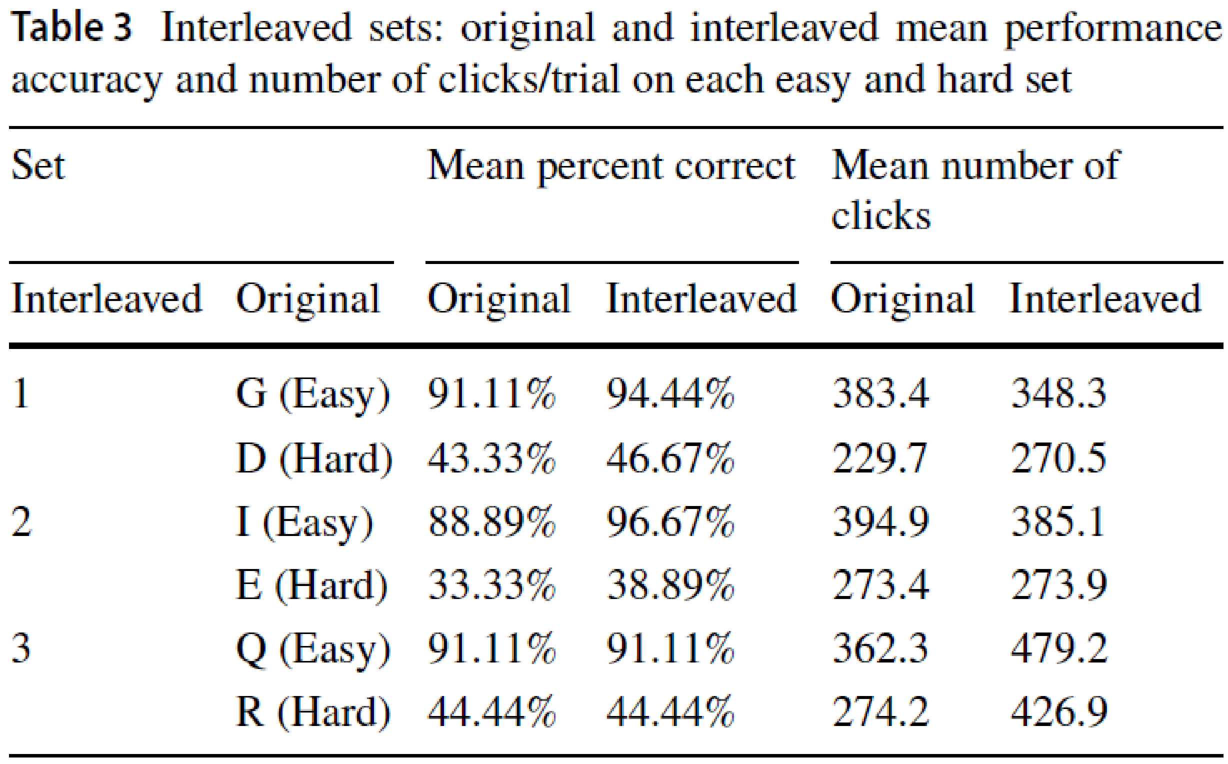


Table 3 Interleaved sets: original and interleaved mean performance accuracy and number of clicks/trial on each easy and hard set

Set		Mean percent correct		Mean number of clicks	
Interleaved	Original	Original	Interleaved	Original	Interleaved
1	G (Easy)	91.11%	94.44%	383.4	348.3
	D (Hard)	43.33%	46.67%	229.7	270.5
2	I (Easy)	88.89%	96.67%	394.9	385.1
	E (Hard)	33.33%	38.89%	273.4	273.9
3	Q (Easy)	91.11%	91.11%	362.3	479.2
	R (Hard)	44.44%	44.44%	274.2	426.9

Evaluating the effects of a change on auditory scenes through "scene shift" sessions

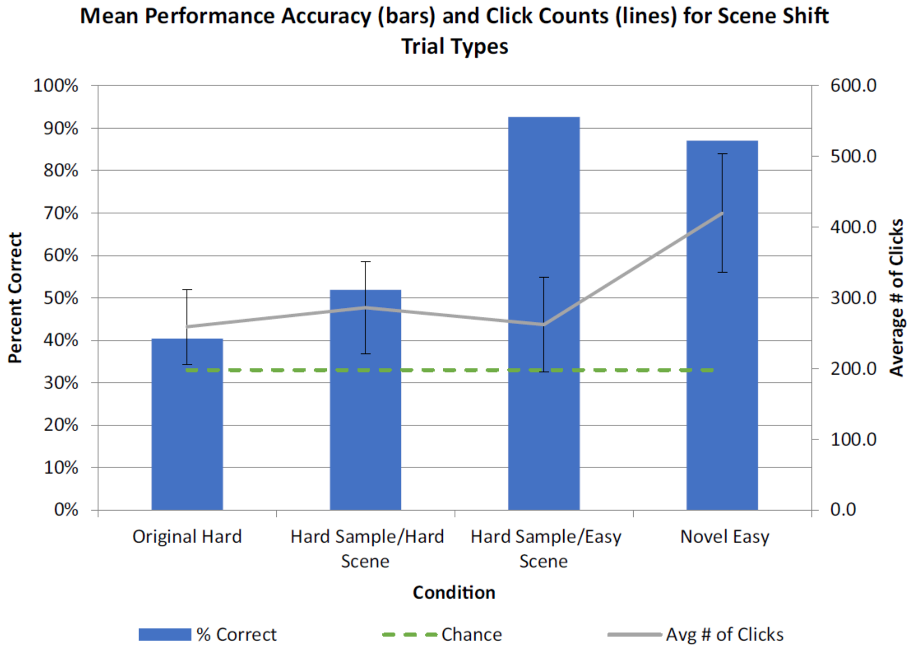
In an early attempt to learn more about how the dolphin's information-seeking behavior would change when familiar objects occurred in new auditory scenes, we tested the dolphin with familiar objects from hard sets (D, E, R) embedded within sessions in which there were 3 trial types: (1) Hard Sample/Hard Scene: 6 trials with the original hard objects presented as usual, (2) Hard Sample/Easy Scene: 6 trials in which the samples were original hard objects and the alternative choice arrays included the hard sample and 2 unfamiliar easy/high-discriminability objects (always 1 PVC object and 1 junk object), and (3) Novel Easy: 6 trials in which the samples were novel easy/high-discriminability objects and the alternative choice arrays included a familiar hard object, an unfamiliar easy PVC object, and an unfamiliar easy junk object. The dolphin experienced 3 sessions of each of these interleaved trials for each of the hard object sets D, E, and R, thus completing 9 sessions total in this condition and resulting in only 18 trials of each trial type/set.

Table 4 presents mean performance accuracy and mean number of clicks/trial for each of the trial types with each of these sets. As expected, the dolphin's performance accuracy, though better than the original accuracy ($\bar{X}$ = 40.37%) with these now familiar objects, was worst on the familiar Hard Sample/Hard Scene trials ($\bar{X}$ = 51.85%), compared to his strong performance on the new objects in the Novel Easy trials ($\bar{X}$ = 87.04%). His performance accuracy with the Hard Sample/Easy Scene trials was also strong ($\bar{X}$ = 92.59%), indicating the power of auditory scene in object recognition. Mean number of clicks/trial to the hard samples remained lower than the easy samples, as previously found, even when the scenes became easy: Original $\bar{X}$ = 259.1 clicks, Hard Sample/Hard Scene $\bar{X}$ = 286.07 clicks, Hard Sample/Easy Scene = 262.2 clicks, Novel Easy = 420.4 clicks. Of course, objects in Novel Easy trials were both highly discriminable and relatively unfamiliar, attributes that resulted in more clicking on previous sets. See Figure 8 for a graphic representation of these results.

Table 4 Scene shifts: performance accuracy and number of clicks/trial for original hard scenes, hard-set samples in hard scenes and unfamiliar easy scenes, and unfamiliar easy scenes

Set	Measure	Original	Hard-set sample		Familiar easy
			Hard scene	Easy scene	
D	Percent correct	43.33%	61.11%	100%	88.89%
	Mean number of clicks	229.7	275.6	285.4	437.1
E	Percent correct	33.33%	44.44%	100%	100%
	Mean number of clicks	273.4	238.8	251.6	435.3
R	Percent correct	44.44%	50.0%	77.78%	72.22%
	Mean number of clicks	274.2	343.8	249.6	388.8

Fig. 8 Performance accuracy (bars) and mean number-of-clicks/trial (line) for scene-shift trial types



Discussion

Our goal in this study was to begin to characterize some of the top-down cognitive processes dolphins might bring to bear for echoic object recognition as targets and auditory scenes that varied in their decipherability methodically shifted from being completely unfamiliar to familiar. We predicted that the dolphin's information-seeking behaviors would change as he learned more about objects and the auditory scenes within which they appeared. The dolphin proved us right.

First, we confirmed that the auditory scenes were differentially decipherable based on the dolphin's performance accuracy with the scenes. Like other dolphins (e.g., DeLong, Au, Harley, & Roitblat, 2007; Xitco & Roitblat, 1996), Calvin could use echolocation to easily discriminate objects in some scenes and barely discriminate objects in others. His highest performance accuracy (100%) for a session was 4.5 times better than his lowest performance accuracy (22%) for a session. We also confirmed that his information-seeking behaviors differed across sessions: The number of clicks he produced to the sample objects varied widely. Again, the highest number of clicks he directed to a sample in a trial (695 clicks) was 4.5 times the lowest number of clicks (155) he directed to a sample in a trial. Of great interest is why.

A comparison of Calvin's investigations of samples related to the four hardest (worst performance accuracy) and four easiest (best performance accuracy) scenes suggested that Calvin's effort, a signal of his attention and motivation, was due to decipherability of scene. For easy and hard object scenes, Calvin reduced his effort as objects became more familiar. However, this reduction was significantly greater for hard scenes versus easy scenes. Investigatory effort changed quickly in the first sessions with the object sets. By the end of the first three trials, Calvin was producing about 64 fewer clicks towards samples in hard object sets versus easy object sets; by the end of the first session, this difference was about 89 clicks. However, in the second and third sessions, he rallied; his effort for easy and hard objects was within 12 and 18 clicks, respectively, for the different scenes. The real plummeting occurred in sessions 4 and 5, with 158 and 138 fewer clicks/trial, respectively, for samples related to hard scenes versus easy scenes. These changes suggest that Calvin was sensitive to the discriminability of the objects in the auditory scenes almost immediately after exposure to them, and he responded by expending less effort to analyze the most difficult scenarios. Then he rebounded by engaging in similar effort no matter the difficulty of the scene, continued to do poorly, and ultimately calibrated his effort based on his experiences with the discriminability of the objects. He continued to put effort into the echoically accessible scenes, though he reduced his effort to some extent as the objects became familiar, but he spent significantly less effort on the impenetrable scenes, similar to Au et al.'s (1982) dolphins in the echoic object detection task when the noise was high. Both dolphins reduced effort, one to the extent that he did not echolocate at all on some trials.

Familiarity does not in itself necessarily lead to reduced effort for object recognition. Because we engaged in a common practice to analyze our data – looking at the extremes (the hardest and easiest scenes) in a complex data set – we also chose to analyze four challenging-but-doable sets to assess more clearly the effect of auditory scene decipherability versus familiarity on effort. In these sets, the dolphin's performance accuracy was low at the start but improved by at least 15% from the first to final sessions. In addition, the dolphin's effort was not significantly lower on the fifth session compared to the first across the sets as a whole, in contrast to the easy and hard scenes. Although there are little other data related to unfamiliar targets and effort in dolphins, we do know that the dolphin Rake continued to work harder to recognize more difficult, versus easier, familiar objects, and that he worked harder to recognize matches when they occurred in different positions.

Calvin's general motivational state does not explain his differential performance accuracy and effort with the hard and easy sets. The fact that Calvin's effort with the hard sets began similarly to the easy sets, that his effort rebounded in sessions 2 and 3 with the hard sets, and that his effort held steady in the challenging-but-doable sets, all suggest that the differences in performance accuracy and effort were not based on vagaries in overall motivation. Nevertheless, we checked this explanation explicitly by interleaving difficult and easy sets all within the same session. In the interleaved sessions, performance accuracy and number of clicks rose to some extent compared to the original sets, perhaps because the objects were more familiar and interleaving allowed a higher reinforcement level in the session given the easy trials, thereby raising motivation; or perhaps because the scenes were changing throughout the session and auditory scene changes may invite more investigation because the dolphin cannot predict which set will be presented. In any case, the object sets remained more (easy sets) and less (hard sets) discriminable in these sessions relative to each other, and the dolphin continued to produce fewer clicks for hard sets versus easy sets. We conclude that some auditory scenes are harder to decipher than others, and the dolphin acts on his abilities to recognize (or not) the objects by reducing his echoic effort when objects are likely to be indiscriminable.

The effect of reinforcement is always a question in studies with trained animals. A few factors suggest that Calvin did not learn to click less with difficult sets versus easy sets just because he received fewer fish for the difficult sets. Firstly, there was a behavioral/time lag between his investigation of the sample (the point at which we measured number of clicks) and his experience of the secondary reinforcer (or its lack) after making a correct choice (or making in incorrect choice). After inspecting the sample, he still had to swim to the alternative array, inspect the alternatives, and make a choice before the secondary bridge whistle was emitted. That is, he performed many behaviors before receiving fish; clicking at the sample was just one of them, and fairly distant from the reinforcers. Secondly, although number of clicks decreased more for hard object sets than for easy object sets, this number decreased for both kinds of sets even though he received a good deal of reinforcement for the easy sets: For easy sets, he received fish for 329 of 360 total trials (i.e., 91% of the trials), and yet he decreased the number of clicks he directed toward the sample. Hence, it is likely not reinforcement that led to his reduced effort, but rather his familiarity with the object sets and the discriminability of the scenes.

Finally, we tried a small number of scene-shift sessions in which objects from difficult scenes were inserted into easy scenes, and though few, these sessions were tantalizing. Within this limited set, the dolphin’s clicks to the hard objects remained fairly steady no matter the scene, even though his performance accuracy in new easy scenes was double his original performance accuracy when these objects were less discriminable. Of course, in this experiment, we gave Calvin very little time to adapt to the new contexts. Within those final Scene Shift sessions, the originally difficult objects only appeared within easy contexts on a third of the trials within each session, and there were only three sessions with each original set for a total of nine sessions. Nevertheless, his stability in effort for hard objects shows that for easily discriminable objects, more clicks are not required for good performance. However, the number of clicks directed towards the unfamiliar easily discriminable objects was 60% higher than those directed to the familiar hard objects. This increased number of clicks is consistent with Calvin’s investigatory behavior throughout the experiment: He often clicked more to unfamiliar objects and to more easily discriminable objects.

Analyzing investigatory movements across scenes

During the scene-shift sessions, Calvin engaged in some interesting information-seeking behaviors that we did not systematically analyze but believe may be important. He would occasionally “peek” at the alternative array by directing his rostrum towards it, while echolocating, on his way to the target sample object. This behavior is of interest because it could have allowed him to assess the auditory scene to come. However, because Calvin did not change his effort when “hard” objects appeared in new easy scenes, his assessment, if completed, did not appear to have a significant effect on his inspection of the sample target. On the other hand, he did well in these new easy scenes without investing more effort, so perhaps he assessed but needed to change nothing in his investigatory effort.

Across the study, object sets were clearly more and less discriminable, and the dolphin responded to that general characteristic. Of great interest is what makes some object sets confusing to dolphins and others not. Some of the object sets varied in material, but many did not. All the hard sets were purely PVC, but so were two of the easy sets and three of the challenging sets. All of the objects varied in shape, and some varied in size. Some were more complex than others. Best would be to collect investigatory clicks and returning echoes from both samples and alternative objects and analyze them in relation to the dolphin's confusions, then use those data to design the next objects with clear hypotheses in mind. Analyzing how a dolphin moves and changes in situations in which she becomes more adept at discrimination would also allow us to determine how the investigation may become more specific – and, hopefully, why.

One investigatory behavior of interest that we did not systematically investigate was Calvin's behavior with target objects at close range. With fair regularity Calvin investigated the sample target object when he was centimeters from it, sometimes even touching the object while buzzing it. These investigations sometimes took the form of outlining the object. Both the behavior and the tight click train could allow the dolphin to gain detailed information given the high repetition rate (Moss & Surlykke, 2001), but what information? And how at such close range? This behavior and its potential functions and outcomes deserve future study.

Conservation

Although this study began very simply in terms of auditory scene analysis, the usefulness of the framework comes in part in considering future work in labs and in the wild. With dolphins, we clearly need future lab work in which we capture and analyze the echoic returns that the dolphin receives, investigations and returns of the auditory scenes/alternative choice arrays themselves, more complex environments, more ecologically relevant targets, eavesdropping scenarios, and more. In the wild, tracking information-seeking behaviors more completely with a specific animal would also improve our knowledge. Taking Moss & Surlykke's (2010) focus on information-seeking behaviors as windows into discerning what kinds of information animals need and how they get it is a powerful framework for helping us think about cognition in echolocating wild cetaceans.

A better understanding of top-down processes, attention, expectations, and motivation in decoding echoic scenes may help us in the conservation of echolocators. For example, allocation of attention is ecologically relevant. If animals engage in less effort in particularly complex auditory scenes, they may miss predators or nets or other important environmental threats. As noted by Malinka et al. (2021), harbor porpoises appear to have the technical capacity to avoid nets, yet they drown in them as bycatch. Could this be an attention problem? Can learning about cues help? Expectations also have a clear effect on the quality of life of dolphins in the wild. For example, Nachtigall and colleagues (2014; 2016) confirmed that warning sounds before loud noises allow animals to prepare for the sound by hearing-dampening to reduce their sensitivity in the expectation that a loud sound will occur. Again, how can we teach them cues to help them prepare for an on-coming din? The investigation of auditory scenes by marine mammals is increasing via work with recording tags and hydrophone arrays (e.g., Ladegaard, Jensen, de Freitas, Ferrera da Silva, & Madsen, 2015; de Freitas, Jensen, Tyne, Bejder, & Madsen, 2015;

Malinka, Tonnesen, Dunn, Claridge, Gridley, Elwen, & Madsen, 2021). Finding a way to incorporate study of top-down processes in these ventures could strengthen our ability to understand and help wild cetaceans. The world they live in is the perceptual world they build, and top-down processes are central to its creation.

Conclusions

Here we found that dolphins engage in top-down processing when working to recognize target objects that appear in auditory scenes that are easier and harder to decipher. They learn about and remember the objects they echolocate and that learning affects their echoic inspections of objects. When objects are unfamiliar, Calvin invests high effort to investigate the objects, and then he responds fairly quickly to his ability to decipher the scenes. With challenging-but-doable scenes, Calvin maintains his effort to decipher the scenes and his object recognition abilities improve with experience. With familiar objects in easily decoded scenes, Calvin produces fewer clicks compared to his original investigations when the objects were unfamiliar, perhaps because he becomes more efficient at recognizing them. With familiar objects in barely decipherable scenes, this pattern is more pronounced: Familiar, difficult scenes receive the lowest relative effort, likely because the dolphin learns the auditory scene is essentially indecipherable in terms of object recognition. Calvin also remembers the objects from difficult scenes and continues to produce a low number of clicks to those “difficult” target objects even when the scenes become more amenable to object recognition, at least in the short term. Overall, memory for particular objects in specific auditory scenes results in a calibration of attention in information-seeking efforts to recognize those objects.

By taking advantage of the public accessibility of information-seeking in a free-swimming echolocating dolphin by counting the clicks he used to investigate target objects that ultimately appeared in simple auditory scenes as the objects moved from being unfamiliar to familiar across 20 different object sets, we were able to learn more about how top-down processing affects echoic investigation of objects in relation to auditory scene. Studies of dolphin echolocation rarely – if ever – include so many unfamiliar objects. Often objects are machined, aspect-independent, and few, sometimes a single “standard”, in order to make it easier for acousticians to predict and decode what is happening with clicks and echoes. For some questions, this approach is powerful. For others, not so much. Dolphins live in complex environments, and somehow they manage to recognize objects echoically in the richness of the world. If we want to understand how they do it, we need to jump into the deep end and figure out what cognitive powers they bring to bear to manage this remarkable feat. That goal means we need to determine what they learn about objects and how that learning affects their investigations of objects within varying scenes. To do that, we need to let them move so they can tell us what they are looking for through their movements, clicks, returning echoes, and dynamic responses to the information they are getting. And to study learning, they have to learn, which means we need to give them a shifting array of unfamiliar objects to study how they adapt as they learn. Our current approach provides a good option for this work if it is tackled by a team of interdisciplinary scientists with expertise in cognitive processes, behavior, acoustics, communication, engineering, neuroscience, modeling, and who are open and flexible collaborators, in order to design, record, analyze, interpret, and model the dolphin’s information-seeking methods. In 1974, Nagel asked, “What is it like to be a bat?” Scientists studying bats and their responses to auditory scenes have made great strides in working to answer that question, a particularly important goal as bat populations around the world face growing environmental threats. Although the barriers of enlisting dolphins into research and the difficulties of working in salt water are formidable, dolphins too are facing substantive environmental threats as ocean temperatures rise, pH levels drop, and coasts face major change due to sea level rise and land subsidence. What will our own effort be in addressing the question, “What is it like to be a dolphin?” Click to answer.

References

Altes, R.A., Dankiewicz, L.A., & Moore, P.W. (2003). Multiecho processing by an echolocating dolphin. *Journal of the Acoustical Society of America*, *114*, 1155-1166. https://doi.org/10.1121/1.1590969

Au, W.W.L., Frankel, A., Helweg, D.A., & Cato, D.H. (2001). Against the humpback whale sonar hypothesis, *IEEE Journal Oceanic Engineering*, *26*, 295–300, Apr. 2001.

Bee, M.A. (2015). Treefrogs as animal models for research on auditory scene analysis and the cocktail party problem. *International Journal of Psychophysiology, 95*, 216-237. http://dx.doi.org/10.1016/j.ijpsycho.2014.01.004

Bilalic, M., McLeod, P., & Gobet, F. (2009). Specialization effect and its influence on memory and problem solving in expert chess players. *Cognitive Science*, *33*, 1117-1143. DOI: 10.1111/j.1551-6709.2009.01030.x

Branstetter, B. K., and Finneran, J. J. (2008). Comodulation masking release in bottlenose dolphins (Tursiops truncatus). *The Journal of the Acoustical Society of America*, *124,* 623–633.

Branstetter, B.K., Finneran, J.J., Fletcher, E.A., Weisman, B.C., & Ridgway, S.H. (2012). Dolphins can maintain vigilant behavior through echolocation for 15 days without interruption or cognitive impairment. *PLoS ONE*, *7*(10): e47478. doi:10.1371/journal.pone.0047478

Branstetter, B.K., Trickey, J.S., Bakhtiari, K., Black, A., Aihara, H., & Finneran, J.J. (2013). Auditory masking patterns in bottlenose dolphins (*Tursiops truncatus*) with natural, anthropogenic, and synthesized noise. *Journal of the Acoustical Society of America*, *133*, 1811-1818. doi: 10.1121/1.4789939

Branstetter, B.K., Van Alstyne, K.R., Strahan, M.G., Tormey, M.N., Wu, T., Breitenstein, R.A., Houser, D.S., Finneran, J.J., & Xitco, M.J. (2020). Spectral cues and temporal integration during cylinder echo discrimination by bottlenose dolphins (*Tursiops truncatus*). Journal of the Acoustical Society of America, 148(2), 614-626. doi: 10.1121/10.0001626

Bregman, A.S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA: MIT Press.

Chiu, C., Xian, W., and Moss, C.F. (2008). Flying in silence: echolocating bats cease vocalizing to avoid sonar jamming. *Proc. Natl. Acad. Sci, 105*, 13115-13120.

Chiu, C., Xian, W., and Moss, C.F. (2009). Adaptive echolocation behavior in bats for the analysis of auditory scenes. *J. Exp. Biol., 212*, 1392-1404.

De Freitas, M., Jensen, F.H., Tyne, J., Bejder, L., & Madsen, P.T. (2015). Echolocation parameters of Australian humpback dolphins (*Sousa sahulensis*) and Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) in the wild. *Journal of the Acoustical Society of America, 137*, 3033-3041. doi: 10.1121/1.4921277

Finneran, J.J., & Branstetter, B.K. (2013). Effects of noise on sound perception in marine mammals. In H. Brumm (ed.), Animal Communication and Noise, Animal Signals and Communication 2, Berlin: Springer -Verlag, pp 273-308. DOI: 10.1007/978-3-642-41494-7_10

Fishman, Y., Micheyl, C., & Steinschneider, M. (2012). Neural mechanisms of rhythmic asking release in monkey primary auditory cortex: Implications for models of auditory scene analysis. *Journal of Neurophysiology, 107(9)*, 2366-2382.

Fishman, Y.I., Steinschneider, M., & Micheyl, C. (2014). Neural representation of concurrent harmonic sounds in monkey primary auditory cortex: Implications for models of auditory scene analysis. *Journal of Neuroscience, 34(37)*, 12425-12443. DOI:10.1523/JNEUROSCI.0025-14.2014

Gotz T, Verfuss, UK, Schnitzler HU (2006) ‘‘Eavesdropping’’ in wild rough-toothed dolphins (Steno bredanensis)? Biol Lett 2:6 –7.

Hayashi, Y. (in press). Molecular mechanism of hippocampal long-term potentiation – Towards multiscale understanding of learning and memory. *Neuroscience Research*. https://doi.org/10.1016/j.neures.2021.08.001

Herman, L. M. (2006). Intelligence and rational behaviour in the bottlenosed dolphin. In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 439–467). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198528272.003.0020

Houser, D. S., and Finneran, J. J. (2006). Variation in the hearing sensitivity of a dolphin population determined through the use of evoked potential audiometry. The Journal of the Acoustical Society of America 120, 4090–4099.

Houser, D., Martin, S.W., Bauer, E.J., Phillips, M. Herrin, T., Cross, M., Vidal, A., & Moore, P.W. (2005). Echolocation characteristics of free-swimming bottlenose dolphins during object detection and identification. Journal of the Acoustical Society of America, 117, 2308-2317.

Itatani, N., & Klump, G.M. (2009). Auditory streaming of amplitude-modulated sounds in the songbird forebrain. *Journal of Neurophysiology, 101(6)*, 3212-3225.

Itatani, N., & Klump, G.M. (2020). Interaction of spatial and non-spatial cues in auditory stream segregation in the European starling. *European Journal of Neuroscience*, *51*, 1191-1200. doi:10.1111/ejn.13716

Janiszewski, C., & Wyer, R.S. Jr. (2014). Content and process priming: A review. Journal of Consumer Psychology, 24(1), 96-118. http://dx.doi.org/10.1016/j.jcps.2013.05.006

Jarvis, J., Bohn, K.M., Tressler, J., & Smotherman, M. (2010). A mechanism for antiphonal echolocation by free-tailed bats. *Animal Behavior, 79*, 787-796.

Kastelein, R.A., Verlaan, M., & Jennings, N. (2008). Number and duration of echolocation clicks trains produced by a harbor porpoise (Phocoena phocoena) in relation to target and performance. *Journal of the Acoustical Society of America, 124,* 40-43. doi: 10.1121/1.2924132

L. N. Frazer and E. Mercado III, “A sonar model for humpback whale song,” IEEE J. Oceanic Eng., vol. 25, pp. 160–182, Jan. 2000

Ladegaard, M., Jensen, F.H., de Freitas, M., Ferrera da Silva, V.M., & Madsen, P.T. (2015). Amazon river dolphins (Inia geoffrensis) use a high-frequency short-range biosonar. Journal of Experimental Biology, 218, 3091-3101. doi:10.1242/jeb.120501

Lewanzik, D., & Goerlitz, H.R. (2021). Task-dependent vocal adjustments to optimize biosonar-based information acquisition. Journal of Experimental Biology, 224, 1-13. doi:10.1242/jeb.234815

Ma, L., Micheyl, C., Yin, P., Oxenham, A.J., & Shamma, S.A. (2010). Behavioral measures of auditory streaming in ferrets (Mustela putorius). *Journal of Comparative Psychology, 124(3)*, 317-330. DOI: 10.1037/a0018273

Malinka, C.E., Rojano-Donate, L., & Madsen, P.T. (2021). Directional biosonar beams allow echolocating harbor porpoises to actively discriminate and intercept closely spaced targets. *Journal of Experimental Biology*, *224,* 1-15. doi:10.1242/jeb.242779

Malinka, C.E., Tonnesen, P., Dunn, C.A., Claridge, D.E., Gridley, T., Elwen, S.H., & Madsen, P.T. (2021). Echolocation click parameters and biosonar behavior of the dward sperm whale (Kogia sima). *Journal of Experimental Biology, 224*, jeb240689. doi:10.1242/jeb.240689

Mercado, E. III, & DeLong, C. M. (2010). Dolphin cognition: Representations and processes in memory and perception. *International Journal of Comparative Psychology, 23*(3), 344–378.

Miller, C.T., & Bee, M.A. (2012). Receiver psychology turns 20: Is it time for a broader approach? Animal Behaviour, 83, 331-343. doi:10.1016/j.tics.2008.02.003

Moore, P., & Finneran, J.J. (2011). Auditory scene analysis in the echolocating dolphin. *Journal of the Acoustical Society of America*, *129*, 2469. https://doi.org/10.1121/1.3588117

Moore, P.W.B., & Pawloski, D.A. (1990). Investigations on the control of echolocation pulses in the dolphin (Tursiops truncatus). In J. Thomas & R. Kastelein (Eds.), *Sensory Abilities of Cetaceans*, New York: Plenum Press, pp. 305-316.

Moss, C.F., & Surlykke, A. (2001). Auditory scene analysis by echolocation in bats. *Journal of the Acoustical Society of America*, *110*, 2207-2226. doi: 10.1121/1.1398051

Moss, C.F., Chiu, C., & Moore, P.W. (2014). Analysis of natural scenes by echolocation in bats and dolphins. In A. Surlykke, et al. (Eds.), *Biosonar*, New York: Springer-Verlag, pp. 231-256. DOI 10.1007/978-1-4614-9146-0_8

Nachtigall, P. E. and Supin, A. Y. (2014). Conditioned hearing sensitivity reduction in a bottlenose dolphin (Tursiops truncatus). *Journal of Experimental Biology*, *217*, 2806-2813.

Nachtigall, P.E., Supin, A.Y., Smith, A.B., & Pacini, A.F. (2016). Expectancy and conditional hearing levels in the bottlenose dolphin (Tursiops truncatus). Journal of Experimental Biology, 219, 844-850. doi:10.1242/jeb.133777

Neilans, E.G., & Dent, M.L. (2015). Temporal coherence for pure tones in budgerigars (Melopsittacus undulates) and humans (Homo sapiens). *Journal of Comparative Psychology, 129(1),* 52-61. http://dx.doi.org/10.1037/a0038368

Pack, A.A., & Herman, L.M. (1995). Sensory integration in the bottlenosed dolphin: Immediate recognition of complex shapes across the senses of echolocation and vision. *Journal of the Acoustical Society of America, 98*, 722-733.

Penner, R.H. (1988). Attention and detection in dolphin echolocation. In P.E. Nachtigall & P.W.B. Moore (Eds.), *Animal Sonar,* New York: Plenum Press, pp. 707-713.

Senzaki, S., Masuda, T., & Ishii, K. (2014). When is perception top-down and when is it not? Culture, narrative, and attention. Cognitive Science, 38, 1493-1506. DOI: 10.1111/cogs.12118

Shinn-Cunningham, B.G. (2008). Object-based auditory and visual attention. Trends in Cognitive Sciences, 12(5), 182-186. doi:10.1016/j.tics.2008.02.003

Shinn-Cunningham, B.G., & Wang, D. (2008). Influences of auditory object formation on phonemic restoration. The Journal of the Acoustical Society of America, 123, 295-301. https://doi.org/10.1121/1.2804701

Shinn-Cunningham, B.G., Lee, A.K.C., & Oxenham, A.J. (2007). A sound element gets lost in perceptual competition. Proceedings of the National Academy of Sciences, 104(29), 12223-12227. www.pnas.org_cgi_doi_10.1073_pnas.0704641104

Smith, P.L., & Little, D.R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic Bulletin and Review, 25*, 2083-2101. https://doi.org/10.3758/s13423-018-1451-8

Smith, P.L., & Little, D.R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic Bulletin and Review, 25*, 2083-2101. https://doi.org/10.3758/s13423-018-1451-8

Stidsholt, L., Johnson, M., Beedholm, K., Jakobsen, L., Kugler, K., Brinklov, S., Salles, A., Moss, C.F., & Madsen, P.T. (2018). A 2.6-g sound and movement tag for studying the acoustic scene and kinematics of echolocating bats. *Methods in Ecology and Evolution, 10*, 48-58. https://doi.org/10.1111/2041-210X.13108

Teilmann, J., Miller, L.A., Kirketerp, T., Kastelein, R.A., Madsen, P.T., Nielsen, B.K., & Au, WWL. (2002). Characteristics of echolocation signals used by a harbor porpoise (*Phocoena phocoena*) in a target detection experiment. *Aquatic Mammals, 28*, 275-284.

Wohlgemuth MJ, Kothari NB, Moss CF (2016) Action Enhances Acoustic Cues for 3-D Target Localization by Echolocating Bats. *PLoS Biology, 14*(9): e1002544. doi:10.1371/journal.pbio.1002544

Xitco M.J., Jr., & Roitblat, H.L. (1996). Object recognition through eavesdropping: passive echolocation in bottlenose dolphins. *Animal Learning Behaviour, 24*(4), 355–365.