

RESEARCH ARTICLE

Visual search in hunting archerfish shares all hallmarks of human performance

Ingo Rischawy and Stefan Schuster*

Universität Bayreuth, Lehrstuhl für Tierphysiologie, D-95440 Bayreuth, Germany

*Author for correspondence (stefan.schuster@uni-bayreuth.de)

SUMMARY

Archerfish are renowned for shooting down aerial prey with water jets, but nothing is known about how they spot prey items in their richly structured mangrove habitats. We trained archerfish to stably assign the categories ‘target’ and ‘background’ to objects solely on the basis of non-motion cues. Unlike many other hunters, archerfish are able to discriminate a target from its background in the complete absence of either self-motion or relative motion parallax cues and without using stored information about the structure of the background. This allowed us to perform matched tests to compare the ways fish and humans scan stationary visual scenes. In humans, visual search is seen as a doorway to cortical mechanisms of how attention is allocated. Fish lack a cortex and we therefore wondered whether archerfish would differ from humans in how they scan a stationary visual scene. Our matched tests failed to disclose any differences in the dependence of response time distributions, a most sensitive indicator of the search mechanism, on number and complexity of background objects. Median and range of response times depended linearly on the number of background objects and the corresponding effective processing time per item increased similarly – approximately fourfold – in both humans and fish when the task was harder. Archerfish, like humans, also systematically scanned the scenery, starting with the closest object. Taken together, benchmark visual search tasks failed to disclose any difference between archerfish – who lack a cortex – and humans.

Key words: visual search, target detection, attention, cortex, mangrove environment.

Received 5 March 2013; Accepted 16 April 2013

INTRODUCTION

Archerfish (*Toxotes* sp.) shoot down prey from overhanging vegetation using a well-aimed shot of water (e.g. Smith, 1936; Lüling, 1963; Schuster, 2007). Interest in these remarkable fish has increased considerably over the past years and now encompasses studies on the shooting mechanisms (e.g. Milburn and Alexander, 1976; Elshoud and Koomen, 1985; Schlegel et al., 2006), the predictive start (e.g. Wöhl and Schuster, 2007; Schlegel and Schuster, 2008; Schuster, 2012), the many outstanding learning capabilities (Schuster et al., 2004; Schuster et al., 2006) and adaptations of their visual processing (e.g. Temple et al., 2010; Ben-Simon et al., 2012). Yet presently we know nothing about how the fish become aware of their potential victims in the first place. A look at the mangrove habitats of these fish readily suggests that this is indeed a demanding problem (Fig. 1). Not only do archerfish have to spot their prey against a richly structured background, but prey items are also surprisingly scarce during daytime when the fish were active in the biotopes we have worked in previously (I.R. and S.S., unpublished). This suggests that these fish should be efficient in quickly spotting a prey item before it takes off again. Moreover, archerfish also must be able to spot a large variety of potential targets without knowing beforehand which types to expect. As opportunistic hunters, they detect and shoot at a wide variety of prey from spiders and insects to small lizards (Smith, 1936) – a property that is mirrored in the way the fish match their maximum force transfer to the scaling of prey adhesive forces (Schlegel et al., 2006).

Earlier experiments (G. Petters and S.S., unpublished) indicate that archerfish, like many other predators, do use prey motion and relative motion parallax cues to detect prey against a structured

background. Here we show that these fish do surprisingly well even in a much harder situation in which the fish are prevented from using any motion cues or stored information about the background objects. This raises the possibility of carrying out matched tests to compare the efficiency of fish and humans in a standard paradigm of human psychophysics: the scanning of stationary flat visual scenes. Ever since the influential papers of Treisman and other pioneers (e.g. Treisman, 1986; Vergheze, 2001; Wolfe, 2010), the field of ‘visual search’ continues to be highly attractive for scientists who view it as the major doorway to understanding how our cortex allocates attention. A typical visual search task consists of a subject locating a target object in an assembly of background (often called ‘distractor’) objects. From the way response time depends on the number of items in the scene and the amount of scrutiny required to discriminate target and background objects, mechanisms have been proposed of how attention is allocated during the search process. In some search tasks the target immediately ‘pops out’ and response time is unaffected by the number of other items present. In a so-called ‘serial search’, median response time increases in proportion to the number of items in the scene. The shape of the distribution of response time and its connection to complexity of the search (e.g. how difficult it is to discriminate between target and background) has recently been found to be a good way to disclose the efficiency and memory capacity of the putative internal tagging of previously scanned objects (Wolfe et al., 2010; Palmer et al., 2011). For instance, a completely amnesic serial search with no internal tagging of already checked non-target objects would produce exponentially distributed response times, and partial tagging or a restricted memory for just a few previously attended items would



Fig. 1. Ecology requires efficient visual search in archerfish. Photos taken from a natural habitat of *Toxotes jaculatrix* and *T. chatareus* in Thailand are shown to illustrate the complexity of the aerial hunting ground of archerfish. To down prey with their renowned shooting behavior, archerfish must first spot suitable prey hidden in the richly structured background of mangrove foliage, twigs and aerial roots. (A) View of mangrove roots with two archerfish crossing. A hunting ground like this can be dry a few hours later due to tidal water movement, forcing the fish to move on. (B) A fish's perspective of the overhanging mangrove foliage. Would you be able to spot the Chrysopidae (green lacewing) underneath one of the mangrove leaves?

translate into response time distributions becoming more skewed (e.g. Palmer et al., 2011).

If the respective mechanisms did indeed depend on a cortex, then matched tests on visual search in non-cortical animals and humans should be reflected in the way response time distributions depend on the task. For instance, response time distributions could be much more skewed and more strongly affected by task complexity. Previous work on bees (e.g. Spaethe et al., 2006; Morawetz and Spaethe, 2012) and birds (e.g. Blough, 1977) has shown that serial search could also be found in these animals. However, no matched tests appear to have been made in animals that would disclose differences in response time distributions. We therefore used the potential opened up by archerfish, an animal that must be efficient in scanning its environment, to test whether hallmarks of visual search – thought to constrain cortical mechanisms – detect differences between fish and humans.

MATERIALS AND METHODS

Fish

Experiments were performed on a group of three adult archerfish [*Toxotes chatareus* (Hamilton 1822)] with a standard length of 12–14 cm. The group was held in a tank of 110×55×50 cm (length×depth×height) filled with brackish water (conductivity: 3.5 mS cm⁻¹) up to a height of 30 cm. Above the aquarium, shielded by a transparent glass plate 35 cm from the water level, an LCD flat screen (22 inch Samsung SyncMaster 2233, Samsung Electronics, Schwalbach am Taunus, Germany) was installed facing down towards the water surface (Fig. 2A). Scenes were presented within a 29 cm diameter circular section (max. visual angle 45 deg). Once the scene was displayed, the first well-directed shot of one of the fish towards the target was considered as a successful location of the target. After each shot, the glass plate was cleared from remaining drops of water to ensure equal visibility in the subsequent trials.

Humans

Each of eight test persons (students of the University of Bayreuth) individually were seated on a chair facing a white wall 135 cm away (from eyes to wall). A video projector was used to create a circular presentation area of 138 cm diameter (max. visual angle 54 deg) right in front of the subjects (Fig. 2B). In order to also require a motor component in the human response time, subjects had to hit the target

with a tennis ball. Unlike the fish, the human subjects were not disturbed by group members but could fully focus on the task. Therefore, we also ran tests in which subjects had to do simple calculations. The calculations were additions and subtractions with numbers from 1 to 100. Subjects were asked to perform one calculation in approximately 3 s. There was no temporal correlation

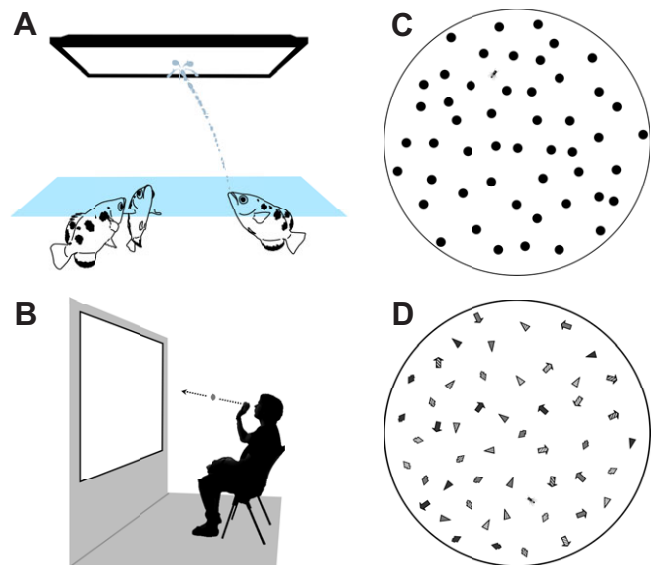


Fig. 2. Matched visual search tasks in archerfish and humans. As we show here, archerfish readily detect non-moving objects in the absence of relative motion parallax. This enables matched tests for both archerfish and human subjects to compare archerfish performance with that of the masters of visual search. (A,B) The presentation area was created using either an LCD screen installed above the aquarium (A) or a video projector (B). A targeted shot (archerfish) or a directed ball (humans) indicated that the subject had spotted the target and selected an appropriate motor response. (C,D) Both archerfish and humans faced identical search tasks: a 'simple' search (C) with the target (the picture of a fly) arranged among identical background items (equally sized black circles) alternated at random with a 'complex' search (D), in which the target was embedded between complex background items differing in shape, orientation and contrast. All sceneries were flat and contained no motion (including parallax) cues, and thus required the subjects to identify the target only by an analysis of non-motion cues.

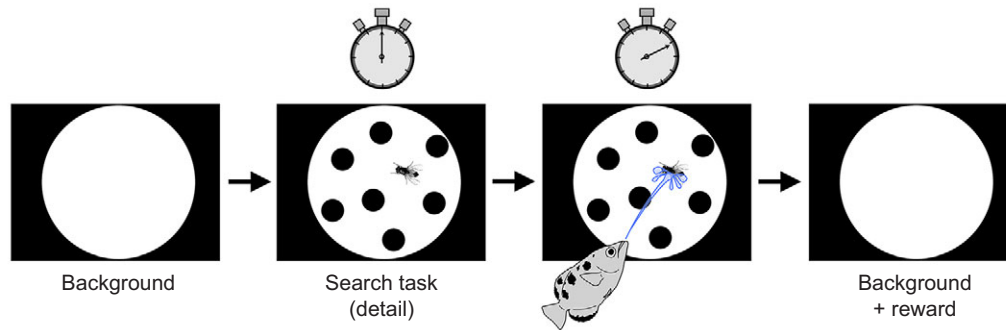


Fig. 3. Schematic to illustrate our visual search tasks. Visual sceneries containing a preselected number of background items plus a target were shown using either an LCD screen (fish) or a video projector (humans). Tests started by showing a white background with no items. Then the search scenery was switched on and timing started. As soon as the target was hit, the time was taken and the scenery switched to the white screen background. Simultaneously with the disappearance of the search task, the fish were rewarded with a dead fly and human subjects received a smile. After a short break, the subsequent search task was presented likewise. Note that the task prevents the use of background memory and relative motion parallax cues.

between the performance of the calculations and the presentation of the stimuli. All subjects were cooperative and readily mastered the calculations at the required rate.

Subjects were treated according the guidelines of the University of Bayreuth and informed consent was obtained from all of them.

Visual scenes, response time and reward

Unless otherwise stated, the following descriptions refer to both archerfish and humans. Subjects were randomly assigned one of 108 visual scenes (see Fig. 2 for examples). These comprised one of nine possible background configurations and – for each of these – 12 pseudo-randomly assigned target locations (with a required minimum distance of 1.31 deg visual angle between objects). Scenes were created in PowerPoint and shown on the LCD flat screen (fish) or projected onto the wall (humans). Within the circular presentation area, the target (the image of a fly) was shown either alone or amidst 25, 50, 75 or 100 background objects. In the ‘simple’ task, all background objects were black dots (Fig. 2C). In the ‘complex’ task, the objects differed in shape and orientation (Fig. 2D). To exclude the possibility that our subjects would somehow remember the location of the background items for each of our nine configurations, we designed two versions of each background that differed only by the locations of the background items on the presentation area. Furthermore, target position was randomized and could be – with equal probability – anywhere within the search area. Both the picture of the fly (target), the black dots (‘simple’ search) and the objects of different shapes (‘complex’ search) were sized 1.0–1.2 cm (1.64–1.96 deg maximum angular extent) on the LCD screen and 5.0–6.0 cm (2.12–2.55 deg maximum angular extent) on the projected area in diameter. Michelson contrast between objects (both target and background objects) and the white background of the scene was 0.84 ± 0.07 (‘fly’), 0.91 ± 0.03 (‘dots’) and 0.63 ± 0.16 (‘shapes’) for the LCD screen and 0.64 ± 0.08 (‘fly’), 0.84 ± 0.005 (‘dots’) and 0.62 ± 0.18 (‘shapes’) for the projected area. Contrast was derived from intensity measurements taken with a precision small-angle intensity meter (Minolta Luminance Meter LS-110, Minolta, Ahrensburg, Schleswig-Holstein, Germany). Generally, experiments started with the circular area being shown without any objects (Fig. 3). The background objects were thus not visible before the target but could be seen only together with the target. Simultaneously with switching on the scene the experimenter started a stopwatch. Upon the first targeted shot fired (fish) or a well-directed ball thrown (human), the experimenter stopped the clock and switched the scene to white again.

To directly measure accuracy and variability of our response time measurements we mimicked the later actual experiments: the experimenter held the stopwatch in one hand and with the other operated the computer keyboard that switched on a visual scene. The scene vanished after a computer-controlled preset time (not known to the experimenter) of 5, 7 or 9 s, which was the signal for the experimenter to stop the watch. This directly gave the measurement-induced latency of 0.28 ± 0.05 s (mean \pm s.d., $N=45$). The inferred variability thus is smaller than the 0.1 s resolution of our stopwatch. Note that the systematic latency (0.28 s) has no relevance for any conclusions in the paper and simply adds to the time it takes the fish to assume the shooting position and to fire. Our finding of serial search allows these unspecific effects to be readily dissociated from those that are specific to the search proper and that can be derived from the slopes of the linear regressions of response time *versus* the number of background items.

After each successful shot, fish were rewarded with a dead fly; humans occasionally received a smile. To reward the fish, immediately after a shot had hit the target, a device fired one dead fly (*Calliphora vicina*, killed by freezing) to a point on the water surface that varied from trial to trial. In the fish, a rewarded task was followed by a pause of at least 30 s. During this time the screen was cleaned and the fish had time to settle and focus on the screen again.

Conventions and statistics

Prior experiments (G. Petters and S.S., unpublished) showed that stable maximum search performance requires archerfish to be kept in at least a small group with intraspecific competition. This required, however, two conventions that were strictly adhered to: (1) no experiment was started when the fish were not swimming calmly below the water surface but instead were chasing each other; and (2) when aggression among group members occurred after the scene was already on, then the task was stopped and no data were taken. All statistics were run using R (version 2.10.1, R Foundation for Statistical Computing, Vienna, Austria). All data were checked for normal distribution by Shapiro–Wilk tests. Data that showed a normal distribution were treated with multivariate linear models; those that did not have a normal distribution were gamma distributed and were treated with linear mixed models. Analyses of the data from human subjects that were either focused or diverted during the search task were treated with a linear mixed model using the identity of the respective subject as a random factor. The significance

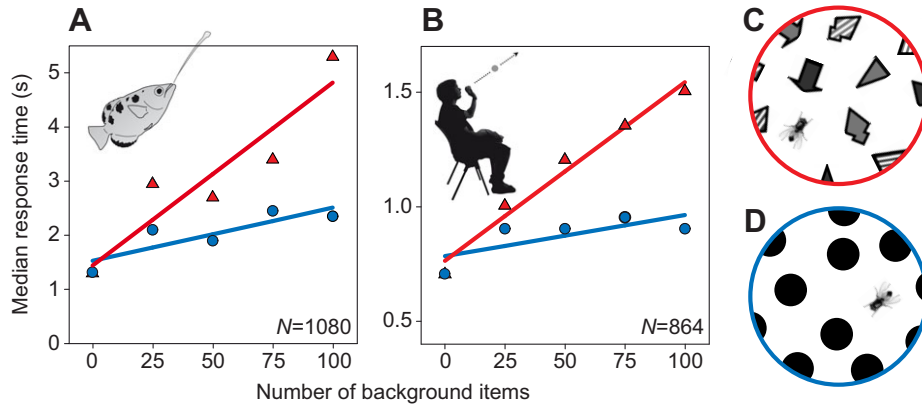


Fig. 4. Humans and archerfish share mechanisms of visual search in the absence of parallax and motion cues. Unlike most other hunting animals, archerfish readily took on our search task in which they could not use self-motion, relative parallax or background memory to spot prey. This allowed us to compare the performance of archerfish with that of humans in a matched visual search task, disclosing surprising similarities. (A,B) Median response time as a function of the number of background items in a scene. Both fish (A) and humans (B) faced either a 'complex' task in which the background items varied in shape, contrast and orientation (C; corresponding data shown in red) or a 'simple' task (D; data in blue). Median response time increased linearly with the number of background items in fish and humans, suggesting an effective processing time per item. While this processing time was faster in humans, it increased similarly in fish and humans (about fourfold) in the 'complex' task, when more scrutiny was needed. Number (N) of responses as indicated.

limit was set at $P=0.05$. In *post hoc* tests, the level of significance was treated by sequential step-down Bonferroni correction (Holm, 1979).

RESULTS

Our experiments started with naive fish that fired at images of a variety of similar-sized targets. From these objects we then selected a variety of shapes plus the image of a fly as items to be shown in the visual search sceneries (see Fig. 2C,D), but we exclusively rewarded shots at the image of the fly. During this phase the fish quickly learned to fire only at the image of the fly and not at any of the other shapes, although these were initially attractive. Training thus had led to an assignment in which the fly was the 'target' and all other previously attractive objects were 'background items'. This assignment was kept throughout the whole study period, as long as we immediately rewarded shots at the fly.

After this initial target-consolidation phase, the very first tests with stationary targets embedded in the same plane as the background already showed that the fish readily spotted the target in complete absence of self-motion or relative motion parallax. Therefore, we abandoned our original plan of training the fish to learn searching without these important cues and started immediately with the tests illustrated in Fig. 3. Note that both background and target appeared simultaneously so that the fish could not detect the target by comparing actual with stored information. The fish readily spotted the non-moving target in the same plane as the background objects and median response time (measured from onset of the presentation until shot fired at target) increased linearly with the number of background objects present in the scenery (linear increase: $P=0.002$; multivariate linear model: $F_{3,6}=15.02$, $P=0.003$, $R^2=0.88$; Fig. 4A). This discovery is a prerequisite that allowed us to settle a problem that would, otherwise, have been difficult to address: response time has a component to it that is independent of the proper search. This comprises the time needed to settle for a shooting position, to aim, to adjust the shooting position to what the other fish do, etc. Our data show that this search-unspecific part of the response time is evident as the offset of the linear relationship between median response time and the number of background items.

The effective processing time per visual item of the scanning mechanism is evident from the slope of the regression line. The slope we find would indicate an effective processing time of 9.8 ms per item (linear regression: $y=0.0098x+1.53$) for the 'simple' task in which all background items were identical. The effective processing time increased significantly ($P=0.034$) to 33.8 ms per item (linear regression: $y=0.0338x+1.44$) in the 'complex' task, in which background items differed so that discriminating them from the target required more scrutiny.

These characteristics were paralleled in our human subjects: response time also increased linearly with the number of background items both for the 'simple' and the 'complex' search (linear increase: $P<0.001$; multivariate linear model: $F_{3,6}=43.38$, $P<0.001$, $R^2=0.96$; Fig. 4B). Effective processing times were 1.8 ms per item in the 'simple' task (linear regression: $y=0.0018x+0.78$) and 7.8 ms per item in the 'complex' task (linear regression: $y=0.0078x+0.76$), respectively, and were thus approximately 5.4 ('simple') and 4.3 ('complex') times shorter in humans than in archerfish. Nevertheless, the relative increase of processing time in the 'complex' task was remarkably similar in fish and humans (3.45 times for fish and 4.34 times for humans). This finding was robust and not attributable to the fact that the human subjects were informed about the task and could fully focus on it. To test this we had the human subjects simultaneously engage in simple calculations while they performed the search task. The added calculations diverted the subjects but affected only the offsets in the plots of response time *versus* background items (linear mixed model: $P>0.001$, $\chi^2=114.25$, d.f.=1) but not the slopes ($P=0.565$, $\chi^2=0.3319$, d.f.=1) – both in the 'simple' and the 'complex' task. Again, with subjects diverted by calculations, the effective processing time per item increased 3.62 times ($P>0.001$, $\chi^2=85.18$, d.f.=1) from the 'simple' to the 'complex' task. This matches the corresponding increase (3.45) in the fish surprisingly well.

So far, probing archerfish in a benchmark visual search task – in which the fish were devoid of motion and parallax cues they would otherwise use – showed no qualitative differences between fish and human performance. A much richer but commonly neglected source of insight into the mechanisms of the search (e.g.

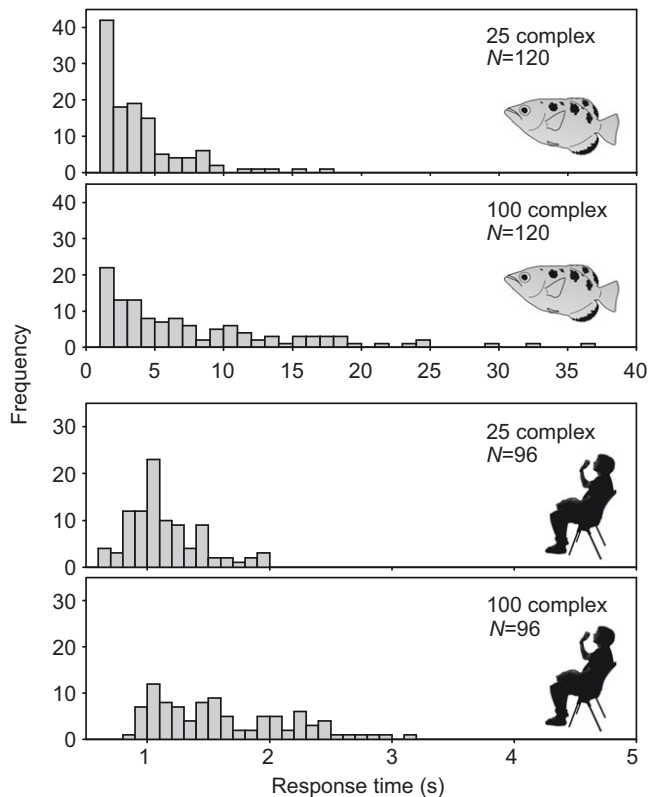


Fig. 5. Examples of response time distributions do not point to differences between archerfish and human. In an attempt to narrow down the underlying search mechanism, we examined in detail the response time distributions for all tests and for both fish and humans. The examples shown here relate to the 'complex' task with either 25 or 100 background items. While detection was faster in the human subjects, the examples do not point to fundamental differences in the distributions: they broaden similarly as the number of background items increase and peaked at low response times. The similarity of the distributions in archerfish and humans is quantitatively analyzed in Fig. 6. Bin width: 1.0 s (fish) and 0.1 s (humans). Number (N) of responses as indicated.

Wolfe et al., 2010) is looking at the shape of the response time distributions and their change with task complexity. Response time distributions were not Gaussian in fish or humans (Shapiro–Wilk test: $P \leq 0.003$; Fig. 5), and in both species broadened linearly with increasing numbers of background items (fish: $P=0.0039$; humans: $P=0.0285$; plots not shown). For a quantitative comparison of response time distributions in humans and fish, we analyzed in detail – and for all search tasks of this account – the two major higher modes of the distributions, skewness and kurtosis (Fig. 6). The analysis provided no overall differences in the distributions in the 'complex' task for fish, humans and humans that simultaneously had to engage in computations (skewness: $P=0.279$, $F_{2,11}=1.44$; kurtosis: $P=0.609$, $F_{2,11}=0.52$; Fig. 6). The only apparent difference between fish and human performance was found in the 'simple' task, in which distributions were more skewed in archerfish than in humans ($P=0.023$). Note, however, that this difference immediately disappeared when the diversion the fish had to face in the group was mimicked in the human subjects by diverting them with the simultaneous calculations ($P=0.41$).

A chance of critically testing our conclusions opened up when one of the fish showed a distinct territoriality and often viewed the scene from the same vantage point. We examined its response times

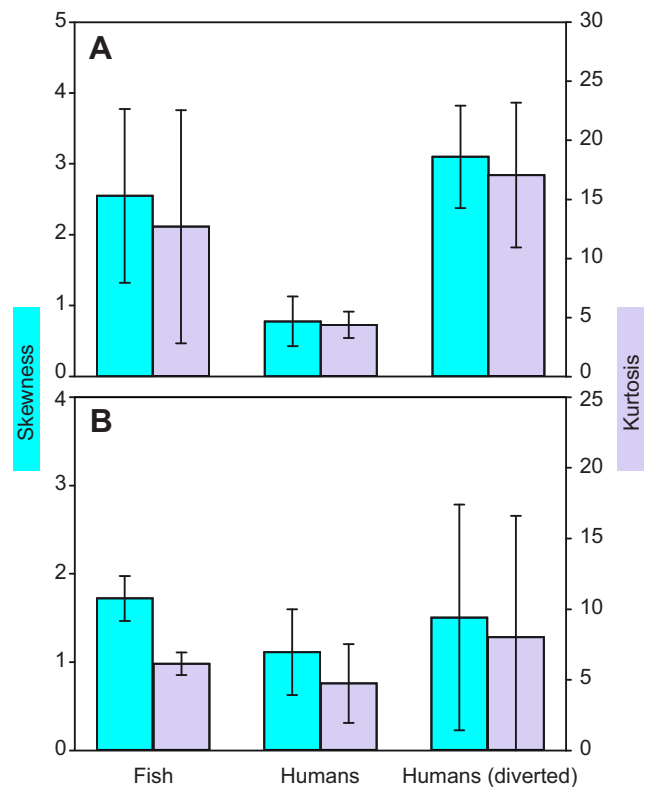


Fig. 6. Hallmark search tasks are unable to provide any evidence for different search mechanisms in humans and archerfish. Results of a detailed statistical analysis of two major modes of the response time distribution, skewness and kurtosis, are shown. (A,B) Mean skewness (left axes) and kurtosis (right axes) in the 'simple' task (A) and the 'complex' task (B) for each of three situations: (1) fish, (2) humans and (3) humans that were diverted by a simultaneous calculation task (see Materials and methods). In the 'simple' task, skewness was significantly higher in fish than in humans without diversion ($P=0.02$), but this apparent difference vanished when human subjects were diverted ($P=0.41$). Kurtosis was not significantly different between fish and humans or fish and diverted humans. In the 'complex' task, no significant differences existed among the three groups both for skewness ($P=0.279$) or kurtosis ($P=0.609$). Data are means \pm s.d. of a total of 2592 responses: 1080 (fish), 756 (humans), 756 (humans diverted).

under such conditions to find out whether effective sampling time depended on where the target was located. In this analysis the circular presentation area (Fig. 7A) was divided in (imaginary) 'proximal' and 'distal' sectors and response time was separately processed depending on whether the target lay in the 'proximal' or in the 'distal' sector. Analyzing the median response times as a function of the number of background items for the two sectors (Fig. 7B) we discovered that the slopes of the regression lines were different (difference in slope: $P=0.002$; multivariate linear model: $F_{3,6}=70.48$, $P>0.001$, $R^2=0.97$), whereas the offsets were not ($P=0.15$). Hence, targets in the 'distal' area are not slowly responded to simply because it took the fish longer to get there and to get ready to fire. Rather, our finding shows that it is indeed the effective processing time per item (and not the offset) that is shorter in the 'proximal' sector and longer in the 'distal' sector – such as if the fish has initially searched the close objects and switches to the distant ones only after it has finished examining all of the closer ones. The findings shown in Fig. 7D support this interpretation. Here, the scenery is divided into 12 (imaginary) sectors (Fig. 7C). With no

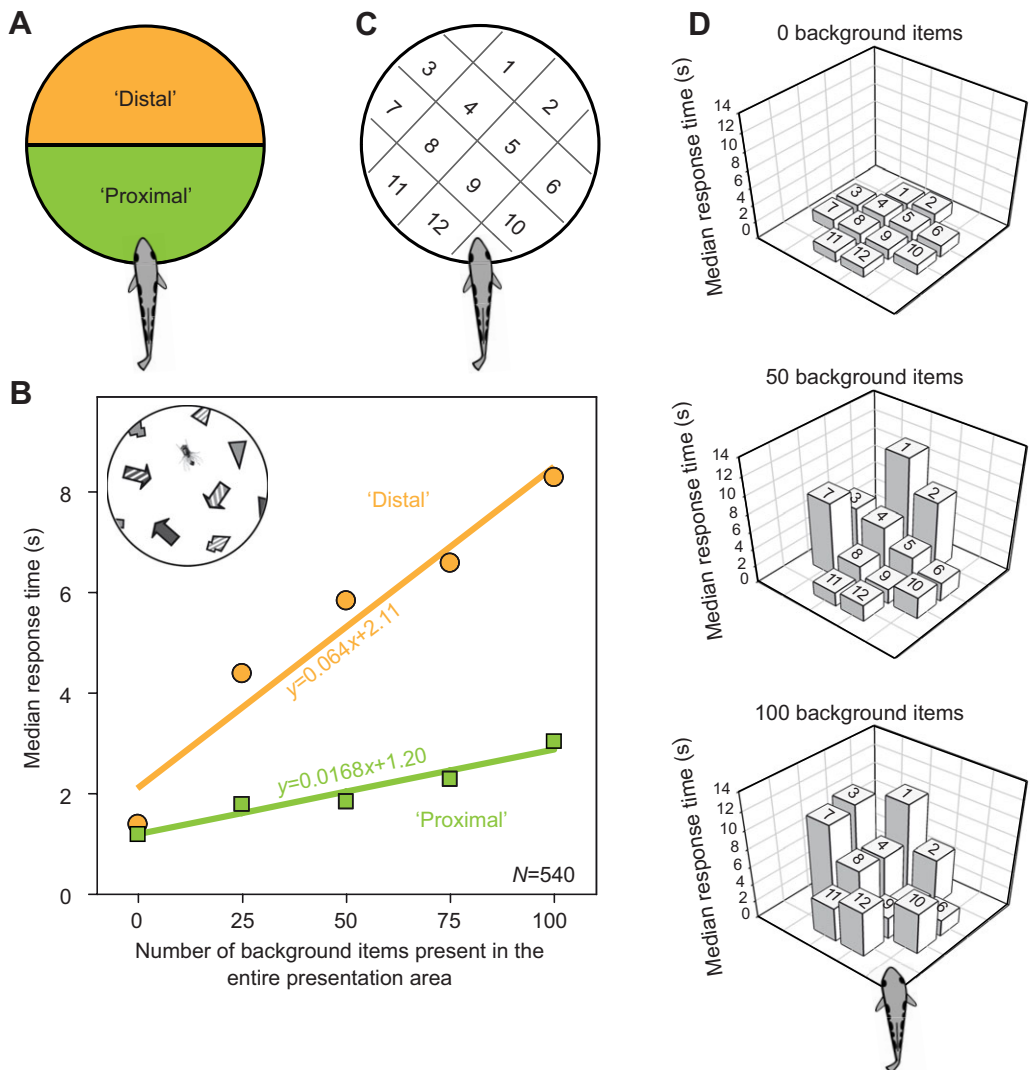


Fig. 7. Archerfish search systematically, starting with the closest spot. (A) Schematic to illustrate how the presentation area was divided into a 'proximal' and a 'distal' sector with respect to the observing fish. (B) Median response time as a function of the number of (complex) background items in the entire presentation area. Inset: aspect of the scenery with target and background items. Note that the slope was significantly ($P=0.002$) higher when the target lay in the 'distal' area than when it was located 'proximally', showing that the difference in response time cannot be attributed to a longer time needed to get ready to shoot when targets were in the 'distal' sector. (C,D) A more detailed view of response time when the scenery (C) was divided into not two but 12 sectors. (D) Response time is reported for each sector when the scenery contained 0, 50 or 100 'complex' background items. Note that the lack of difference in response time when only the target was present ('0 background items') is consistent with the finding shown in B of no significant connection between getting ready to fire and distance to target. In the same manner, if background items are present, a target in the close sector is spotted considerably faster.

background items present, median response times were independent of target location. When background items were present, then response times were always short when the target appeared close to the fish and longer when the target lay in the distant parts of the scene.

DISCUSSION

The major surprise of this study is that hunting archerfish can scan a flat visual scenery based solely on non-motion cues and do this in ways that benchmark tests cannot discriminate from human performance. In both species, median response times but also the range of response times increased linearly with the number of background items in a scene. When more scrutiny was needed to discriminate target and background items, the effective processing time per item increased in surprisingly similar manner in both fish and humans. Furthermore, a detailed analysis of the higher moments of the response time distributions – a powerful tool to analyze memory for scanned objects (Palmer et al., 2011) – failed to show any distinct difference between the way archerfish and humans scanned the scenes.

Comparing archerfish and human performance

Comparing absolute performance levels among animals is tricky and not often as profitable as it seems. Our study was designed

to compare functional relationships between fish and humans, but not to report tasks (such as that shown in Fig. 1B) in which archerfish would certainly fare much better than humans. If one did compare the absolute performance levels, then our study would seem to imply that humans scanned approximately 4.3–5.4 times faster than fish. This comparison would already account for the differences in the search-unspecific response time: getting ready and showing the required motor response was different in fish and humans, but this could be dissected out from the way response time depended on the number of background objects (Fig. 4). Nevertheless, it is still not profitable to compare the absolute levels of performance: in contrast to the human subjects, fish moved around freely and had to judge the scenery from all possible orientations and distances. In experiments that run over longer periods, it is important to keep the fish in groups (G. Petters and S.S., unpublished), which causes differences in how much fish and humans could focus on the task. Our attempt to divert the human subjects by having them simultaneously make calculations reduced the amount of focus for the human subjects, but it would be rather naive to claim that this distraction was in any way quantitatively matched with that of the fish. Many more points could be raised, but most importantly, archerfish were tested in a challenging situation in which we had prevented them from using cues they would otherwise use.

Nevertheless, focusing on functional relationships clearly showed that the existing diagnostic tools, including some whose importance has only recently been stressed (e.g. Wolfe et al., 2010; Palmer et al., 2011), failed to detect any difference in the mechanisms that archerfish and humans employed in scanning our stationary scenes: (1) median response time increased linearly with the number of background items (Fig. 4); (2) the effective scan time per item increased in the same proportion when the task required more scrutiny (Fig. 4; 'simple' versus 'complex'); and (3) no difference could be spotted in either the shape of the response time distributions or in the way they depended on the number of background items and task complexity (Figs 5, 6).

Is the 'serial search' of fish and humans serial?

Ever since Treisman (e.g. Treisman, 1986), a linear increase of median response time has often been interpreted as indicating that (1) the internal search proceeds serially, scanning object by object until the target is detected and that (2) each object is internally scanned only once. From these assertions, it is evident that median response time increases in proportion to the number of objects that need to be scanned: with N objects that need to be internally classified (each in time τ) as background items or targets, the average total time needed is $\tau N/2$. However, most authors do not seem to be interested in the second conclusion that also follows from the assertions: the response time distribution would have to be flat with a range that also increases linearly with the number of objects. A look at the distributions (Fig. 5) shows that response times were not uniformly distributed in fish or humans. This indicates that interpretations 1 and 2 are far too simple. Probing further into differences of the search requires a detailed look at the behavior of the response time distributions (Figs 5, 6). This analysis also failed to detect any qualitative difference between humans and fish, thus supporting the notion that both species scan stationary scenes at least with computationally similar algorithms.

Our findings thus show that both species deviate from the standard view of serial searching. But what are they scanning? The findings shown in Fig. 7 suggest a starting point for such an analysis, using trained fish. In Fig. 7B, the effective per item processing time increased approximately fourfold when the target lay in the distal sector. This is difficult to explain if scanning proceeds item wise. But it is easy to explain if subareas were scanned. Depending on the assumptions made on memory of which subsets had already been scanned, a rough calculation suggests that these subareas could be surprisingly large, but more evidence would be needed to speculate any further.

Complex ecological demands may be the basis for the efficiency of visual search in archerfish

The remarkable capability of archerfish to efficiently search a target in the complete absence of motion or motion parallax cues appears to be rather rare among predators. This ability and its efficient use are probably linked with the high demands of searching prey in a complex mangrove environment. The fish have to spot a variety of prey animals, some even well camouflaged, from various distances within the richly structured aerial background of their habitat. Moreover, the environment does not allow fixed hunting territories in which the fish could potentially memorize the visual background. In their natural habitats, the interaction of the tides with freshwater inflow (I.R. and S.S., unpublished) makes fluctuation of water levels difficult to predict – with two major consequences: first, a suitable hunting ground cannot be kept (because it will become dry); and second, when

leaving the area, it is unknown when the spot can be used again. Our finding that the fish did so well without being able to memorize the background is probably related to this – the fish could not have evolved simple 'novelty' mechanisms in which they stored the aerial background of their 'hunting territory' and detected any deviations from the stored memory templates. Because there is no simple territory that the fish can memorize and because prey are rare, it is very likely that the fish will not be looking when a prey item is landing. This could be one reason why archerfish had to develop efficient ways to spot non-moving prey items.

Certainly, many other animals may share efficient search mechanisms with humans. Aspects of serial search have, for instance, been discovered in honeybees (Spaethe et al., 2006; Morawetz and Spaethe, 2012), whose lifestyle also makes them excellent candidates for highly efficient search with remarkable memory for rejected non-target items. A comparative approach, particularly on animals with small brains or animals that can employ only small parts of their brains during the task, will help us discover the constraints on neural circuitry for efficient search.

Conclusions

Our findings suggest that demands such as those that archerfish face in their mangrove habitats can cause even a fish brain to implement mechanisms that in humans, and presumably other mammals, are linked to their cortex. Our findings raise doubt that visual search data can constrain cortical architectures based on findings of response time distributions and the way response times depend on the number of items in a scenery. In fact, as we show here, these factors cannot even discriminate humans from an animal that completely lacks a cortex. Obviously, the need to efficiently find objects has not entered the world with the advent of cortices. Our findings thus support the rather natural view that many animals must have come up with algorithms that are similarly effective to those used by humans and that these mechanisms may not depend on a cortex. Studying such animals could help discover more general network constraints for efficient search.

ACKNOWLEDGEMENTS

We thank Drs Machnik and Schulze for valuable discussions, Antje Halwas and Karl-Heinz Pöhner for technical support, Katja Keller and Michaela Hahn for help in testing the human subjects, and Dr Stefan Gross for superb statistical guidance.

AUTHOR CONTRIBUTIONS

S.S. and I.R. conceived the project, I.R. performed the experiments, and S.S. and I.R. analyzed the data and wrote the paper.

COMPETING INTERESTS

No competing interests declared.

FUNDING

Supported by grants of the Deutsche Forschungsgemeinschaft (SCHU1470/2, 7 and 8).

REFERENCES

- Ben-Simon, A., Ben-Shahar, O., Vasserman, G., Ben-Tov, M. and Segev, R. (2012). Visual acuity in the archerfish: behavior, anatomy, and neurophysiology. *J. Vis.* **12**, 1-19.
- Blough, D. S. (1977). Visual search in the pigeon: hunt and peck method. *Science* **196**, 1013-1014.
- Elshoud, G. C. A. and Koomen, P. (1985). A biomechanical analysis of spitting in archer fishes. *Zoomorph.* **105**, 240-252.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **6**, 65-70.
- Lüling, K. H. (1963). The archerfish. *Sci. Am.* **209**, 100-108.

- Milburn, O. and Alexander, R. M.** (1976). The performance of the muscles involved in spitting by the archerfish *Toxotes*. *J. Zool.* **180**, 243-251.
- Morawetz, L. and Spaethe, J.** (2012). Visual attention in a complex search task differs between honeybees and bumblebees. *J. Exp. Biol.* **215**, 2515-2523.
- Palmer, E. M., Horowitz, T. S., Torralba, A. and Wolfe, J. M.** (2011). What are the shapes of response time distributions in visual search? *J. Exp. Psychol. Hum. Percept. Perform.* **37**, 58-71.
- Schlegel, T. and Schuster, S.** (2008). Small circuits for large tasks: high-speed decision-making in archerfish. *Science* **319**, 104-106.
- Schlegel, T., Schmid, C. J. and Schuster, S.** (2006). Archerfish shots are evolutionarily matched to prey adhesion. *Curr. Biol.* **16**, R836-R837.
- Schuster, S.** (2007). Archerfish. *Curr. Biol.* **17**, R494-R495.
- Schuster, S.** (2012). Fast-starts in hunting fish: decision-making in small networks of identified neurons. *Curr. Opin. Neurobiol.* **22**, 279-284.
- Schuster, S., Rossel, S., Schmidtman, A., Jäger, I. and Poralla, J.** (2004). Archer fish learn to compensate for complex optical distortions to determine the absolute size of their aerial prey. *Curr. Biol.* **14**, 1565-1568.
- Schuster, S., Wöhl, S., Griebisch, M. and Klostermeier, I.** (2006). Animal cognition: how archer fish learn to down rapidly moving targets. *Curr. Biol.* **16**, 378-383.
- Smith, H. M.** (1936). The archer fish. *Nat. Hist.* **38**, 2-11.
- Spaethe, J., Tautz, J. and Chittka, L.** (2006). Do honeybees detect colour targets using serial or parallel visual search? *J. Exp. Biol.* **209**, 987-993.
- Temple, S., Hart, N. S., Marshall, N. J. and Collin, S. P.** (2010). A spitting image: specializations in archerfish eyes for vision at the interface between air and water. *Proc. Biol. Sci.* **277**, 2607-2615.
- Treisman, A.** (1986). Features and objects in visual processing. *Sci. Am.* **254**, 114-125.
- Vergheze, P.** (2001). Visual search and attention: a signal detection theory approach. *Neuron* **31**, 523-535.
- Wöhl, S. and Schuster, S.** (2007). The predictive start of hunting archer fish: a flexible and precise motor pattern performed with the kinematics of an escape C-start. *J. Exp. Biol.* **210**, 311-324.
- Wolfe, J. M.** (2010). Visual search. *Curr. Biol.* **20**, R346-R349.
- Wolfe, J. M., Palmer, E. M. and Horowitz, T. S.** (2010). Reaction time distributions constrain models of visual search. *Vision Res.* **50**, 1304-1311.