

Category-specific attention for animals reflects ancestral priorities, not expertise

Joshua New^{*†‡}, Leda Cosmides^{*}, and John Tooby^{*}

^{*}Center for Evolutionary Psychology, University of California, Santa Barbara, CA 93106; and [†]Department of Psychology, Yale University, New Haven, CT 06520

Edited by Gordon H. Orians, University of Washington, Seattle, WA, and approved August 10, 2007 (received for review May 3, 2007)

Visual attention mechanisms are known to select information to process based on current goals, personal relevance, and lower-level features. Here we present evidence that human visual attention also includes a high-level category-specialized system that monitors animals in an ongoing manner. Exposed to alternations between complex natural scenes and duplicates with a single change (a change-detection paradigm), subjects are substantially faster and more accurate at detecting changes in animals relative to changes in all tested categories of inanimate objects, even vehicles, which they have been trained for years to monitor for sudden life-or-death changes in trajectory. This animate monitoring bias could not be accounted for by differences in lower-level visual characteristics, how interesting the target objects were, experience, or expertise, implicating mechanisms that evolved to direct attention differentially to objects by virtue of their membership in ancestrally important categories, regardless of their current utility.

animacy | category specificity | domain specificity | evolutionary psychology | visual attention

Visual attention is an umbrella term for the set of operations that select some portions of a scene, rather than others, for more extensive processing. These operations evolved because some categories of information in the visual environment were likely to be more important or time-sensitive than others for activities that contributed to an organism's survival or reproduction. The selection criteria that direct visual attention can be categorized by their origin: (i) goal-derived: criteria activated volitionally in response to a transient internally represented goal; (ii) ancestrally derived: criteria so generally useful for a species, generation after generation, that natural selection favored mechanisms that cause them to develop in a species-typical manner; and (iii) expertise-derived: criteria extracted during ontogeny by evolved mechanisms specialized for detecting which perceptual cues predict information that enhances task performance.

These three types of criteria may also interact; for example, differential experience or temporary goals could calibrate or elaborate ancestrally derived criteria built into the attentional architecture.

The ways in which human attention can be affected by goals and expertise have been extensively investigated. Indeed, humans are zoologically unique in the extent to which we evolved to engage in behavior tailored to achieve situation-specific goals as a regular part of our subsistence and sociality (1, 2). Among our foraging ancestors, improvising solutions in response to the distinctive features of situations would have benefited from the existence of goal-driven voluntary attentional mechanisms. As predicted by such a view, otherwise arbitrary but task-relevant objects command more attention than task-irrelevant ones (3), and expertise in a task domain shifts attention to more task-significant objects (4), features (5), and locations (6).

In contrast, attentional selection criteria that evolved in response to the payoffs inherent in the structure of the ancestral world have been less systematically explored. Yet, the rapid identification of the semantic category to which an object

belongs (e.g., animal, plant, person, tool, terrain) and what its presence in the scene signifies [e.g., predatory danger, food (prey), offspring at risk] would have been central to solving many ancestral adaptive problems. That is, stably and cumulatively across hundreds of thousands of generations, attention allocated to different semantic categories would have returned different average informational payoffs. From this perspective, it would be odd to find that attention to objects was designed to be deployed in a category-neutral way. Yet there has been comparatively little research into whether some semantic categories spontaneously recruit more attention than others, and whether such recruitment might be based on evolved prioritization. Most exceptions have studied attention and responses to highly social information such as faces (7, 8), eye gaze (9), hand gestures (10), and stylized human outlines (stick drawings and silhouettes) (11).

The Animate Monitoring Hypothesis

For ancestral hunter-gatherers immersed in a rich biotic environment, non-human and human animals would have been the two most consequential time-sensitive categories to monitor on an ongoing basis (12). As family, friends, potential mates, and adversaries, humans afforded social opportunities and dangers. Information about non-human animals was also of critical importance to our foraging ancestors. Non-human animals were predators on humans; food when they strayed close enough to be worth pursuing; dangers when surprised or threatened by virtue of their venom, horns, claws, mass, strength, or propensity to charge; or sources of information about other animals or plants that were hidden or occluded; etc. Not only were animals (human and non-human) vital features of the visual environment, but they change their status far more frequently than plants, artifacts, or features of the terrain. Animals can change their minds, behavior, trajectory, or location in a fraction of a second, making their frequent reinspection as indispensable as their initial detection.

For these reasons, we hypothesized that the human attention system evolved to reliably develop certain category-specific selection criteria, including a set designed to differentially monitor animals and humans. These should cause stronger spontaneous recruitment of attention to humans and to non-human animals than to objects drawn from less time-sensitive or vital categories (e.g., plants, mountains, artifacts). We call this the animate monitoring hypothesis. Animate monitoring algorithms are hypothesized to have coevolved alongside goal-driven

Author contributions: J.N., L.C., and J.T. designed research; J.N. performed research; J.N. and L.C. analyzed data; and J.N., L.C., and J.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: CD, change detection; Exp *n*, Experiment *n*; RT, reaction time.

See Commentary on page 16396.

[†]To whom correspondence should be addressed. E-mail: joshua.new@yale.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0703913104/DC1.

© 2007 by The National Academy of Sciences of the USA

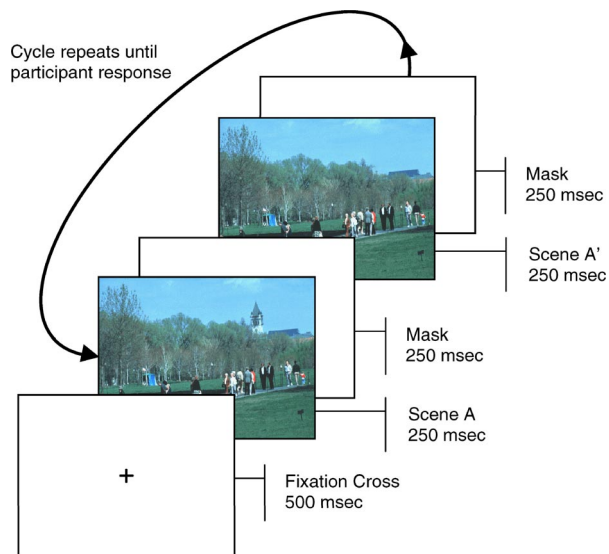


Fig. 1. Diagram illustrating the sequence and timing of each trial in Exp 1–5.

voluntary processes that focus attention on task-relevant objects, providing the voluntary system with one of several interrupt circuits made necessary by a surprising world. These algorithms should operate automatically and autonomously from executive function, so that important changes in non-humans and humans can be detected rapidly, even when they are unexpected or irrelevant to current goals or activities. Hence, we propose that animate inputs will recruit visual attention in a way that is less context-, goal-, expertise-, and state-dependent than other inputs. Although increasingly focused attention may increasingly screen out task-irrelevant stimuli, such exclusion should affect human and animal stimuli less than members of other categories. In particular, subjects' attention should display the predicted animate monitoring bias in the absence of instructions to look for animals or humans and regardless of their relevance to the task or to subjects' goals.

The counterhypothesis is that visual attention contains no mechanisms designed to differentially allocate attention on the basis of the semantic category of the input. This means there should be no mechanisms that evolved to deploy attention differentially to animate targets, and therefore no animate monitoring bias should be found. If, nevertheless, evidence of such a bias were to be found, the fallback hypothesis would be that such an effect would be the result of expertise: that is, starting with an equipotential attentional system, ontogenetic training would accrete attentional biases as a function of differential experience with the stimulus inputs and their ontogenetic importance. We will call this the expertise hypothesis.

Assessing Preferential Attention

Experiments show that viewers often fail to detect sizeable changes in an image when these occur during very brief interruptions of sight, a phenomenon known as change blindness (13, 14). To explore the selection criteria implemented by attentional mechanisms, we used the change detection (CD) paradigm (Fig. 1), in which viewers are asked to spot the difference between two rapidly alternating scenes that are identical except for a change to one object. The logic is straightforward: in a CD paradigm, changes to more attended objects or regions in a complex natural scene will be detected faster and more reliably than changes to less-attended ones. By varying which features in a scene are changed, one can learn the criteria by which visual attention mechanisms select objects for further processing. In a CD

experiment, subjects are instructed to detect changes, but they are not given any task-specific goal that would direct their attention to some kinds of objects over others. Thus, the CD paradigm can be used to investigate how attention is deployed in the absence of a voluntary goal-directed search (15). If the animate bias hypothesis is correct, then change blindness will be attenuated for animals and humans compared with other object categories. This is because category-specific attention mechanisms will automatically check the status of animals and people on an ongoing basis.

We adapted a standard CD task (14) to test for the predicted category-specific biases (Fig. 1). The stimuli were color photographs of natural complex scenes (Fig. 2). For Experiments (Exp) 1–4, 70 scenes with target objects from five semantic categories were used (14 in each category): two animate (people and animals) and three inanimate [plants; moveable/manipulable artifacts designed for interaction with human hands/body (e.g., stapler, wheelbarrow); fixed artifacts constructable as topographical landmarks (e.g., windmill, house)]. These categories were chosen because converging evidence from neuropsychology and cognitive development suggests each is associated with a functionally distinct neural substrate (16, 17). Each involves an evolutionarily important category, but only the animates require close visual monitoring. Target categories for Exp 5 (96 scenes) were vehicles, artifacts that do not move on their own, non-human animals, and people. [For details, see supporting information (SI) Appendix 1].

Tests and Predictions

If, as hypothesized, the human attentional architecture includes evolved mechanisms designed to differentially direct attention to both human and non-human animals, then, in a CD task using complex natural scenes, we predict that: (i) changes to animals (both human and non-human) will be detected more quickly than changes to inanimate objects and (ii) changes to animals will be detected more frequently than changes to inanimate objects. By hypothesis, attention is differentially recruited to animals by virtue of neural processes recognizing (at some level) their category membership. The bias is category-driven. Therefore, (iii) although animals will be judged more interesting than inanimate objects, detection rates will be better predicted by the target's category (animate or inanimate) than by how interesting the targets are judged to be, and (iv) the detection advantage for animate categories will not be due to lower-level perceptual characteristics, such as visual complexity or high contrast.

According to the expertise counterhypothesis, any effects by category will arise from differences in frequency of observation, differential training with different categories, or the relative importance during ontogeny of paying differential attention by category. We selected vehicles as an evolutionarily novel contrast category with which subjects have a great deal of experience; which move and do so in a self-propelled fashion; and which subjects have been trained from childhood as pedestrians and drivers to differentially attend to because of the life-or-death importance of anticipating their momentary shifts in trajectory. In comparison, our subjects see and interact with non-human animals far less often than with vehicles, and animals have little practical significance to our subjects. Despite greater subject expertise with vehicles, we predict that (v) the animate bias will not be a consequence of ontogenetic exposure to things in motion. In particular, although subjects see large numbers of vehicles moving every day, changes to vehicles will not be detected as quickly or reliably as changes to animals and people.

Finally, this study affords an opportunity to measure the effects of expertise on visual attention. Subjects have a lifetime of intensive training in attending to one species above all: humans. In contrast, subjects have orders-of-magnitude less experience attending to any other given species. The difference

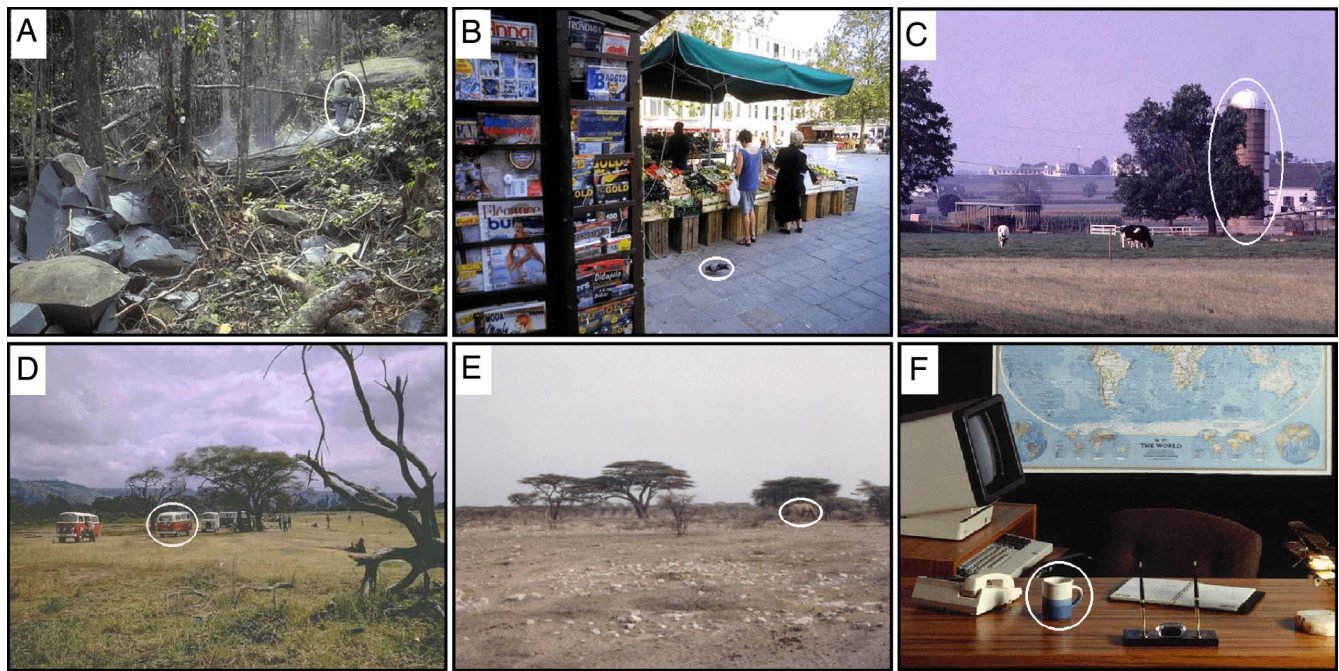


Fig. 2. Sample stimuli with targets circled. Although they are small (measured in pixels), peripheral, and blend into the background, the human (A) and elephant (E) were detected 100% of the time, and the hit rate for the tiny pigeon (B) was 91%. In contrast, average hit rates were 76% for the silo (C) and 67% for the high-contrast mug in the foreground (F), yet both are substantially larger in pixels than the elephant and pigeon. The simple comparison between the elephant and the minivan (D) is equally instructive. They occur in a similar visual background, yet changes to the high-contrast red minivan were detected only 72% of the time (compared with the smaller low-contrast elephant's 100% detection rate).

in performance between attention to humans and attention to other animal species gives a measure of the importance of expertise in training attention to animate inputs.

Results

Exp 1 was designed to test predictions *i–iii* (above) of the animate bias hypothesis, Exp 2 was a replication of Exp 1, and Exp 3–5 were designed to test predictions *iv–v*.

The hit rate (percent correct) was used to assess accuracy, because false alarms were so rare across the five experiments (2% of all responses; *SI Appendix 1.1*). Reaction times (RTs) are for hits.

Do Animals and People Recruit Preferential Attention? Yes. Changes to animals and people were detected more often and more quickly than changes to inanimate objects in Exp 1 and 2 (Fig. 3 *A* and *B*). More specifically, changes to animate targets (animals and people) were detected faster than changes to inanimate ones (plants, moveable artifacts, and fixed artifacts), both in Exp 1 and its replication (Exp 2); animate vs. inanimate target RTs: $P = 10^{-10}$ and 10^{-15} , respectively. Changes to animate targets were detected 1–2 seconds faster than changes to inanimate ones, and the effect size (r) associated with this difference was large in both experiments (0.88 and 0.86).

The greater speed in detecting changes to animals and people was not achieved at the expense of accuracy. On the contrary, subjects were faster and more accurate for animate targets, which elicited hit rates 21–25% points higher than inanimate targets (Exp 1 and 2, $r = 0.84$ and 0.80 ; $P = 10^{-8}$ and 10^{-10} ; false-alarm rates were low, 0.92% and 1.6%). Overall, 89% of changes to animate targets were correctly detected vs. 66% of changes to inanimate ones. The animate advantage in speed and accuracy remains strong, even when inanimates are compared only to non-human animals (see Fig. 3; RT, $r = 0.80$ and 0.64 , $P = 10^{-7}$ and 10^{-11} ; hits, $r = 0.82$ and 0.63 , $P \leq 0.0002$).

Following convention, we reported RTs for hits only. However, this measure fails to capture cases in which a change to the target was missed entirely; missing a change is a more severe case of “change blindness” than being slow to notice one. Subjects were change-blind more often for inanimate targets than for animate ones (miss rates, 34% inanimate vs. 11% animate). Because this is not reflected in mean RTs for hits, the difference between animate and inanimate RTs underestimates the animate attentional advantage. Moreover, mean RTs can mask important differences in the time course of change detection.

Fig. 3 addresses these concerns by showing, for each category, the time course of change detection. The relationship between time elapsed and total number of changes detected is plotted. Steeper slopes indicate earlier change detection; higher asymptotes mean more changes were eventually detected (i.e., less change blindness). Consistent with the hypothesis that animals and people should undergo incidental monitoring so that changes in their location and state can be rapidly detected, the curves for the two animate categories have steeper slopes and asymptote at higher levels than those for the three inanimate categories. Moreover, there appear to be attentional capture as well as monitoring effects.

Attentional Capture. The animate and inanimate curves diverge quickly: there were more hits for animate than for inanimate targets even for the fastest responses, ones in which changes were detected in <1 second (Exp 1, hits 8.8% vs. 3.9%; $P = 0.0025$, $r = 0.52$, no false alarms; Exp 2, hits, 3.8% vs. 1.6%, $P = 0.002$, $r = 0.48$; one false alarm). This suggests that animates capture attention in addition to eliciting more frequent monitoring. The maximal difference between animate and inanimate detection occurred at 3.5–4 elapsed seconds, a 33–37% point difference, with an effect size of $r > 0.93$ (P values = 10^{-14}).

Do Animals and People Receive Preferential Attention Because They Are More “Interesting?” In CD studies, interesting items are detected faster than uninteresting ones (14, 18). When a separate

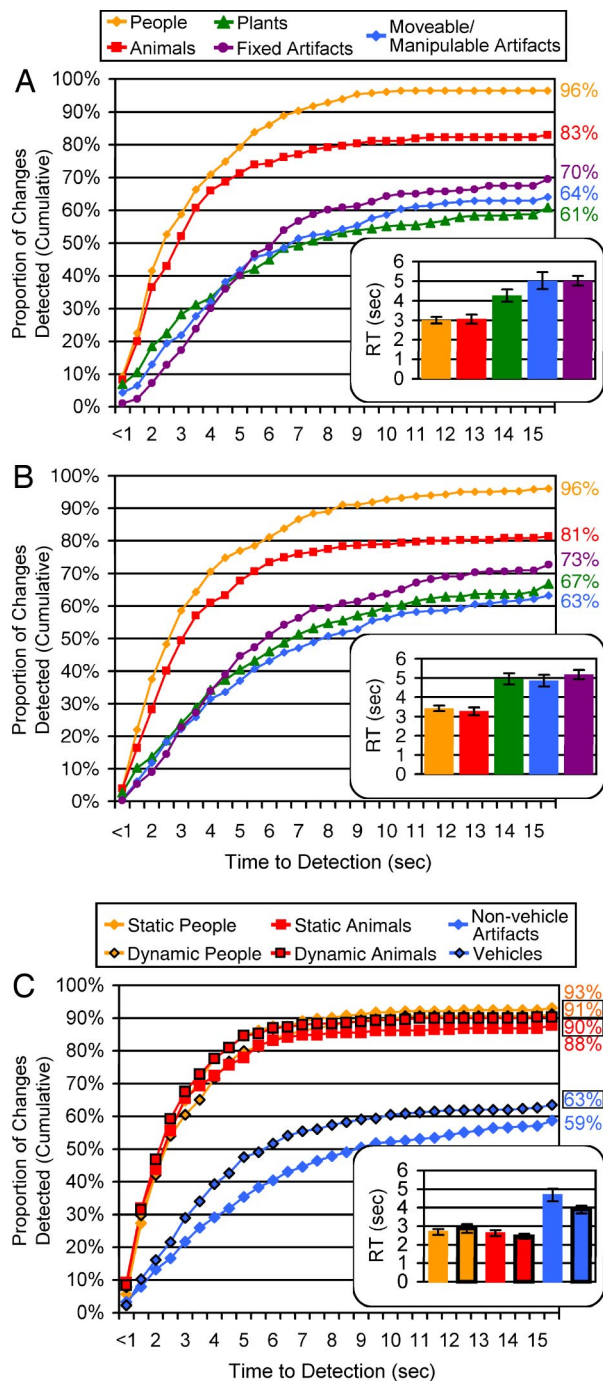


Fig. 3. Changes to animals and people are detected faster and more accurately than changes to plants and artifacts. Graphs show proportion of changes detected as a function of time and semantic category. (Inset) Mean RT for each category (people, animals, plants, moveable/manipulable artifacts, and fixed artifacts). (A) Results for Exp 1. Animate targets: RT $M = 3,034$ msec (SD, 882), hit rate $M = 89.8\%$ (SD, 7.4). Inanimate targets: RT $M = 4,772$ msec (SD, 1,404), hit rate $M = 64.9\%$ (SD, 15.7). (B) Results for Exp 2. Animate targets: RT $M = 3,346$ (SD, 893), hit rate $M = 88.7\%$ (SD, 8.0). Inanimate RT $M = 4,996$ (SD, 1,284), hit rate $M = 67.5\%$ (SD, 16.5). (C) Results for Exp 5. RT: animate $M = 2,661$ msec (SD, 770). Hit rate, animate vs. vehicle: 90.6% (SD, 7.8) vs. 63.5% (SD, 18.8), $P = 10^{-15}$.

group of subjects rated how interesting each target was (*SI Appendix 1*), interest ratings correlated with animacy ($r = 0.60$, $P = 10^{-7}$). But does this explain the animate attention effect?

No. Although they were correlated with animacy, interest

ratings do not predict RTs once one statistically controls for whether the target is animate (partial $r = -0.16$, $P = 0.20$). In contrast, animate targets elicit faster RTs, even after controlling for interest ratings (partial $r = -0.41$, $P = 0.001$; see *SI Appendix 1.2*). The same result holds for hit rates (animacy, partial $r = 0.37$, $P = 0.002$; interest, partial $r = 0.064$, $P = 0.60$). Thus, the animacy bias was not a side effect of animals and people being more “interesting” targets, as judged by deliberative processes. Fast, accurate change detection of animates results from a category-driven process: animacy, not interest, predicts change detection efficiency.

Is Preferential Attention to Animates a Side Effect of Differences in Lower-Level Visual Characteristics? Does the animate attention effect found in Exp 1 and 2 reflect nothing more than a confound in the stimuli? Given that lower-level features (e.g., color, luminance) can affect attention in simple visual arrays (19) and more natural and complex scenes (20), it is important to eliminate the hypothesis that, in the particular stimuli we used, these features differed for animate vs. inanimate targets.

Target luminance, size (pixels), and eccentricity were entered into a multiple regression model; none predicted RT or accuracy, either across or within domains (P values range from 0.2 to 0.9). To eliminate counterhypotheses invoking any potential lower-level feature(s), Exp 3 and 4 were conducted.

Inverting photos perfectly preserves their lower-level stimulus properties but makes identifying the semantic category to which a target belongs more difficult (8, 18, 21). Further, scene inversion sizably reduces the detection of changes to high-relative to low-interest items (18) (but see ref. 22). If lower-level properties are causing the animate attention advantage, then it should appear even when photos are inverted. In contrast, if the attentional bias is category-driven, then manipulations that interfere with categorization but preserve lower-level perceptual features should eliminate the animate change-detection advantage.

Exp 3 was identical to Exp 1 and 2, except the photos were inverted. The procedure and stimuli for Exp 4 were also the same, except target category identification was disrupted not by inverting but by blurring each photo with a Gaussian blurring function in Adobe Photoshop (see *SI Appendix 2*). This preserves many lower-level characteristics (although not as perfectly as inversion) but disrupts object recognition more than inversion does. Both manipulations were used, because each has advantages that the other lacks. Each method succeeded in disrupting recognition; compared with Exp 1 and 2, RTs were slower in Exp 3 and 4, and accuracy was worse overall in Exp 4 (*SI Appendix 1.3*). If lower-level characteristics were causing the animate attention effect, then it should still appear in Exp 3 and 4. It did not (see Fig. 4).

Specifically, inverting scenes eliminated the animate advantage in detection speed ($P = 0.25$). Changes to inverted people, animals, fixed artifacts, and plants elicited comparable mean detection times (Fig. 4A; *SI Appendix 1.4*). When inverted, accuracy was comparable for fixed artifacts, plants, people, and animals. (Compared with other inanimate targets, accuracy for inverted moveable artifacts was disproportionately low, a pattern not seen in the upright conditions; *SI Appendix 1.5*). This is in contrast to the pattern for upright scenes, where animate beings showed a consistent speed and accuracy advantage compared with all inanimate categories.

Blurring upright scenes also eliminated the animate advantage in detection speed ($P = 0.17$). There was no animate advantage in accuracy either. In fact, the reverse was true: in the blur condition, accuracy was greater for inanimate objects (Fig. 4B; *SI Appendix 1.6*).

Inversion and blurring disrupt recognition, which is necessary for targets to be identified as animate vs. inanimate, while

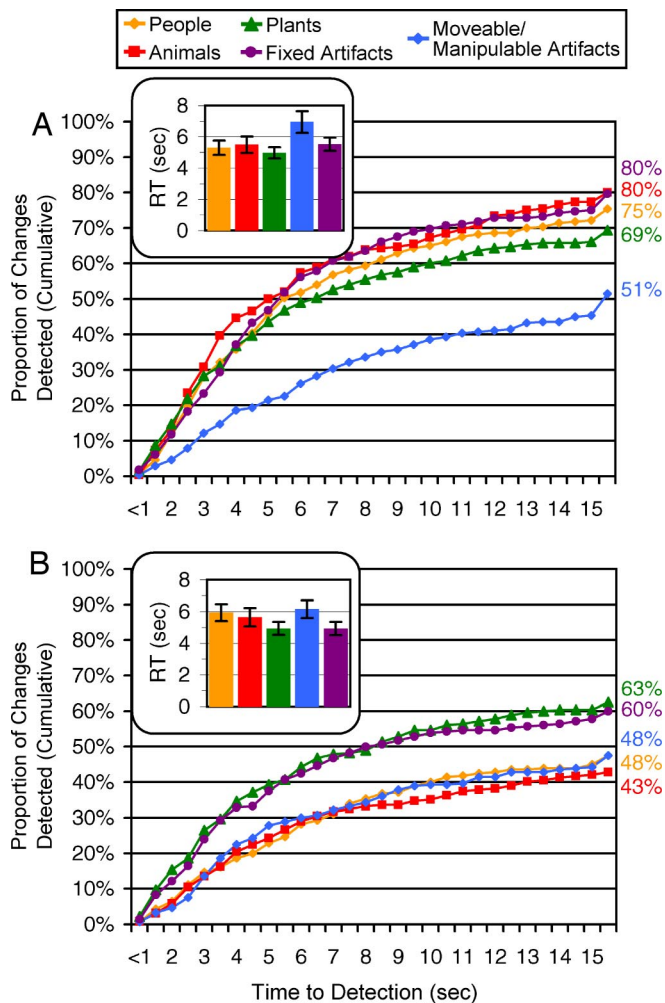


Fig. 4. Disrupting recognition eliminates the advantage of animates in change detection, showing that the animate advantage is driven by category, not by lower-level visual features. Graphs show proportion of changes detected as a function of time and category when recognition is disrupted. (Inset) Mean RT for each category. (A) Results for Exp 3 using inverted stimuli. RT, animate $M = 5,399$ (SD, 2,139), inanimate $M = 5,813$ (SD, 2,405). (See [SI Appendices 1.4 and 1.5](#).) (B) Exp 4, blurred stimuli. RT, animate $M = 5,792$ (SD, 2,705), inanimate $M = 5,337$ (SD, 2,121). Accuracy; animate $M = 45.2\%$ (SD, 15.1), inanimate $M = 56.7\%$ (SD, 13.5), greater accuracy for inanimates; $P = 0.0001$, $r = 0.67$.

preserving all (inversion) or some (blurring) lower-level stimulus characteristics. That these two manipulations eliminated the animate detection advantage found in Exp 1 and 2 indicates the animate attentional advantage depends on recognition of the target's semantic category. It was not a side effect of incidental differences in contrast, visual complexity, or any other lower-level property of the stimuli used. (Additional controls show that the animate advantage also remains strong when controlling for potential differences in scene backgrounds; see [SI Appendix 1.7](#)).

Is Preferential Attention to Animates a Consequence of Experience with Motion? The animate monitoring hypothesis proposes that animates are attended by virtue of category-specific attentional mechanisms that are triggered by properties of animals and humans, not by mechanisms that attend to anything often seen in motion. Vehicles were chosen as a control category, because they move yet are not animals.

Vehicles are seen in motion every day, and the failure to monitor that motion has life-or-death consequences. Indeed, this

expertise might give vehicles a detection advantage over other inanimate objects. But the prediction that animate inputs will be closely attended is not based on expertise-driven attention. It is based on the hypothesis that the visual system is designed to monitor animates because of their importance ancestrally. Consequently, animals and people should be monitored more closely than vehicles, despite our ontogenetic experience of vehicular motion and its importance to survival in the modern world. To test this, we conducted Exp 5, which specifically compared detection of animate targets to vehicles.

Of the artifact targets, 24 were vehicles (on roads, rivers, etc.), and 24 were artifacts that do not move on their own (e.g., lampposts, keys). To see whether there is any effect of implied motion on attention (23–25) due not to the target's category but rather to representations of motion or momentum (e.g., sitting vs. walking), half the people and half the animals were in motion, and half were not. Thus, there were static and dynamic animate targets and static (lampposts, keys) and dynamic (vehicles) inanimate targets. Otherwise, the procedure was the same as for Exp 1.

The results of Exp 5 are shown in Fig. 3C. Accuracy for vehicles and static artifacts was low (and comparable), with changes to vehicles detected faster than changes to static artifacts ($P = 0.00072$, $r = 0.52$). Nevertheless, changes to animals and people were detected >1 second faster than changes to vehicles, and the effect size was large, $r = 0.82$ (animate vs. vehicles, $P = 10^{-11}$). Even so, this underestimates the animate attentional advantage over vehicles, because accuracy for animate targets was 27% points higher than for vehicles, another large effect size, $r = 0.87$ [animate vs. vehicle, 90.6% (SD, 7.8) vs. 63.5% (SD, 18.8), $P = 10^{-12}$]. That is, subjects were change blind >36% of the time for vehicles but <10% of the time for animals and people. Detection of animate targets was better than vehicle targets at all time intervals, even <1 second.

Compared with vehicles, the speed and accuracy advantage for non-human animals was just as large as for people (animals vs. vehicles, RT, $r = 0.80$, $P = 10^{-10}$; hits, $r = 0.84$, $P = 10^{-14}$; people vs. vehicles, RT, $r = 0.78$, $P = 10^{-9}$; hits, $r = 0.88$, $P = 10^{-16}$). Moreover, the advantage for non-human animals remains just as large if the vehicle category is restricted to include only cars and trucks, the vehicles that subjects need to monitor most often (RT, $r = 0.79$, $P = 10^{-9}$; hits: $r = 0.85$, $P = 10^{-15}$).

To make sure these effects were not due to incidental differences in low-level visual characteristics, we conducted an inversion control for Exp 5 (analogous to Exp 3). Although there were some differences between categories on inversion, the animate attentional advantage in Exp 5 remains large and significant when these potential differences in low-level features are controlled for (RT, $r = 0.74$, $P = 10^{-7}$; hits, $r = 0.88$, $P = 10^{-12}$; [SI Appendix 1.8](#)). The same is true when one controls for potential differences in scene backgrounds ([SI Appendix 1.7](#)).

It is known that the human visual system has a bias to detect motion, and that momentum is represented even from still pictures (23–25). Are changes to animals and people detected faster and more accurately merely as a side effect of attention to objects in motion, animate or not?

No. For the animate monitoring effect to be a side effect of motion detection, there would have to be a CD advantage for targets in motion over stationary ones, even for the categories animal and person. Fig. 3C shows this was not the case; for animals and people, CD was just as fast and accurate when their pose was stationary as when it was dynamic (stationary vs. dynamic; hit RT means 2,660 msec (SD, 968) vs. 2,661 (SD, 1,142); hit rates, 91% for both). Thus implied motion does not cause a category-independent increase in attentional monitoring.

Because there were no category-independent effects of representational momentum on change detection, such effects cannot explain the CD advantage of vehicles over static artifacts.

This suggests that the vehicle vs. static advantage was caused by greater monitoring of objects identified as vehicles (whether in motion or not).

Better change detection for non-human animals than for vehicles demonstrates a category-based dissociation between recognition and monitoring. In directed categorization tasks, the visual system can rapidly detect the presence of both animals and vehicles in a natural scene (26), even in the near absence of attention (27). But the large difference in change detection demonstrated here shows that the attentional system spontaneously monitors animals more than vehicles (or other inanimates), even when there is no instruction to do so.

Is There an Effect of Ontogenetic Expertise? The CD advantage of vehicles over other inanimate objects is consistent with a modest expertise effect, although it could also be a side effect of an animate attention bias that is weakly evoked by vehicles (people make vehicles move; psychophysically, vehicular motion exhibits the contingent reactivity of animate motion) (28). But if experience were having a major effect on incidental attention, we would see a large CD advantage for vehicles over animals. Instead, the reverse is true. There would also be a large CD advantage for humans over non-human animals, a prediction that is also falsified.

In modern environments, encounters with other humans are more frequent and have greater consequences than encounters with non-human animals. So how much (or little) does ontogenetic expertise with humans promote change detection, compared with non-human animals? The curves for animals and humans are almost identical in Exp 5 (Fig. 3C), and they track each other closely for time intervals <3–4 seconds in Exp 1 and its replication (Exp 2). More specifically, in Exp 1, 2, and 5, there was no speed advantage for humans over animals (animals vs. humans, mean RT for hits, $P = 0.83, 0.46, 0.07$; animals were faster). Accuracy was the same in Exp 5 ($P = 0.07$) but higher for humans than for animals in Exp 1 and its replication (Exp 1, $P = 0.0003, r = 0.61$; Exp 2, $P = 10^{-7}, r = 0.76$).

Close attention to non-human animals makes sense in ancestral environments but not in the ontogenetic environment experienced by our subjects. Moreover, subjects are visually trained on the human species many orders of magnitude more than on any other species. If expertise acquisition was a function of frequency of exposure and stimulus importance, then change detection for human targets should be orders of magnitude better than for non-human animal targets. Yet there was no speed advantage for detecting changes to humans, and a lifetime of exposure to humans led only to an inconsistent advantage in accuracy: more changes were detected when the target was a person than an animal in Exp 1 and its replication but not in Exp 5. The limited differences in outcome compared with massive

differences in training indicate that other causes are at play aside from, or in addition to, training. These results, like the animal–vehicle difference, call into serious question ontogenetic explanations that invoke only domain-general expertise learning.

Conclusion

Changes to animals, whether human or non-human, were detected more quickly and reliably than changes to vehicles, buildings, plants, or tools. Better change detection for non-human animals than for vehicles reveals a monitoring system better tuned to ancestral than to modern priorities. The ability to quickly detect changes in the state and location of vehicles on the highway has life-or-death consequences and is a highly trained ability; indeed, driving provides training with feedback, a situation that should promote the development of expertise-derived selection criteria. Yet subjects were better at detecting changes to non-human animals, an ability that had life-or-death consequences for our hunter–gatherer ancestors but is merely a distraction in modern cities and suburbs. This speaks to the origin of the selection criteria that created the animate monitoring bias.

The selection criteria responsible were not goal-derived: the only instructed goal was to detect changes (of any kind), and there was nothing in the structure of the task to make animals more task-relevant than inanimate objects (if anything, the reverse was true: there were more changes to inanimates than to animates). Nor were they expertise-derived: in the modern world, detecting changes in animals is an inconsequential and untrained ability compared with detecting changes in vehicles. Taken together, the results herein implicate a visual monitoring system equipped with ancestrally derived animal-specific selection criteria. This domain-specific subsystem within visual attention appears well designed for solving an ancient adaptive problem: detecting the presence of human and non-human animals and monitoring them for changes in their state and location.

Materials and Methods

Five CD experiments were conducted, each involving a different set of subjects (*SI Appendix 1*). The 70 scenes used in Exp 1–4 are shown in *SI Appendices 3–7*. The 96 scenes used in Exp 5 are shown in *SI Appendices 8–11*; there were 48 with artifact targets and 48 with animate targets (24 people and 24 animals). Of the artifact targets, 24 were vehicles, and 24 were artifacts that do not move on their own.

We thank Max Krasnow, Christina Larson, June Betancourt, Jack Loomis, and Howard Waldow and appreciate the support of the University of California, Santa Barbara (UCSB) Academic Senate (L.C.); UCSB Graduate Division (J.N.); National Institute of Mental Health National Research Service Award (F32 MH076495-02, to J.N.); and National Institutes of Health Director's Pioneer Award (to L.C.).

- Cosmides L, Tooby J (2000) in *Metarepresentations: A Multidisciplinary Perspective*, ed Sperber D (Oxford Univ Press, New York), pp 53–115.
- Tooby J, DeVore I (1987) in *Primate Models of Hominid Behavior*, ed Kinzey W (SUNY Press, New York), pp 183–237.
- Shinoda H, Hayhoe M, Shrivastava A (2001) *Vision Res* 41:3535–3545.
- Werner S, Thies B (2000) *Visual Cognit* 7:163–173.
- Myles-Worsley M, Johnston W, Simons M (1988) *J Exp Psychol Learn Mem Cognit* 14:553–557.
- Chun M, Jiang Y (1998) *Cognit Psychol* 36:28–71.
- Mack A, Rock I (1998) *Inattentional Blindness* (MIT Press, Cambridge, MA).
- Ro T, Russell C, Lavie N (2001) *Psychol Sci* 12:94–99.
- Friesen C, Kingstone A (1998) *Psychon Bull Rev* 5:490–495.
- Langton S, Bruce V (2000) *J Exp Hum Percept Perform* 26:747–757.
- Downing P, Bray D, Rogers J, Childs C (2004) *Cognition* 93:B27–B38.
- Orians G, Heerwagen J (1992) in *The Adapted Mind*, eds Barkow J, Cosmides L, Tooby J (Oxford Univ Press, New York), pp 555–579.
- Grimes J (1996) in *Vancouver Studies in Cognitive Science*, ed Akins K (Oxford Univ Press, New York), Vol 5, pp 89–110.
- Rensink R, O'Regan J, Clark J (1997) *Psychol Sci* 8:368–373.
- Shapiro K (2000) *Visual Cognit* 7:83–91.
- Caramazza A (2000) in *The New Cognitive Neurosciences*, ed Gazzaniga M (MIT Press, Cambridge, MA), pp 1199–1210.
- Caramazza A, Shelton J (1998) *J Cognit Neurosci* 10:1–34.
- Kelley T, Chun M, Chua K (2003) *J Vision* 2:1–5.
- Turatto M, Galfano G (2002) *Vision Res* 40:1639–1644.
- Parkhurst D, Law K, Niebur E (2002) *Vision Res* 42:107–123.
- Rock I (1974) *Sci Am* 230:78–85.
- Shore D, Klein R (2000) *J Gen Psychol* 127:27–43.
- Freyd J (1983) *Percept Psychophys* 33:575–581.
- Kourtzi Z, Kanwisher N (2000) *J Cognit Neurosci* 12:48–55.
- Senior C, Barnes J, Giampietro V, Simmons A, Bullmore E, Brammer M, David A (2000) *Curr Biol* 10:16–22.
- Van Rullen R, Thorpe S (2001) *Perception* 30:655–668.
- Li F, VanRullen R, Koch C, Perona P (2002) *Proc Natl Acad Sci USA* 99:9596–9601.
- Blakemore S, Boyer P, Pachot-Clouard M, Meltzoff A, Segetbarth C, Decety J (2003) *Cereb Cortex* 13:837–844.