Cognition in the Fast Lane: Ravens' Gazes are Half as Short as Humans' when Choosing Objects

Katarzyna Bobrowicz*[1] and Mathias Osvath[1]

[1] Department of Philosophy and Cognitive Science, Lund University, Sweden

ORCID 0000-0002-3469-6339 (KB), ORCID 0000-0002-7873-0930 (MO)

*Corresponding author (Email: katarzyna.bobrowicz@lucs.lu.se)

Abstract –Time cannot be directly perceived; instead, its flow is inferred from the influx of sensory information. To prevent sensory overload, attentional mechanisms split up information into processable units. This portioning remains imperceptible to the individual. However, the length of these units still influences the speed of perception and the speed at which behaviors are performed. Previous studies have focused on establishing the length of these units in various mammalian species – mainly humans – by measuring different types of behaviors, including gaze. However, no such studies have been conducted on birds. We measured duration of ravens' (*Corvus corax*) single gazes towards selectable objects before a choice was made, and compared it with that of humans in a similar set up. The raven gaze durations were approximately half those of humans (which fell slightly short of previously established ranges). We hypothesize that these differences are mainly due to the much higher so-called flicker-fusion-frequency in birds, which makes their vision faster in the sense that it picks up more information per time unit than mammalian vision does. We further discuss that the speed of perception might influence the general speed of cognitive processing in more complex tasks as well, and suggest that the addition of a temporal component in comparative cognitive studies might be informative.

Keywords: Time constant, Raven, Moment, Temporal processing, Cognition, Visual perception

As time escapes direct perception, its flow is inferred from the stream of information reaching the perceptual systems (Gibson, 1975). Continuous and rapid influx would soon overload these systems and cripple responses to the surrounding world; therefore, the attentional mechanisms pitch in and chop the information into smaller portions for further processing (Pöppel, 2004). Ultimately, these portions blend into a continuous experience, and the segmentation vanishes; like when separately shot frames are put together and deliver a smooth motion picture.

Following this analogy, the attentional mechanisms impose a certain frame rate on perception. Although unnoticeable in individual experience, they also impose a fixed frame rate on the consequent sensorimotor responses. This frame rate, controlled by underlying neurobiological processes, differs between species (Baer, 1862; Gerstner & Goldberg, 1994; Schleidt & Kien, 1997), which can be observed in frame-by-frame analyses of videotaped behaviors. To estimate the duration of such frames, previous studies have investigated a range of behaviors, from hugs (humans: Pöppel, 2004) and sneezing (non-primate mammals: Gerstner & Goldberg, 1994), through hammering and thatching, to making butter and

lighting a fire (humans: Schleidt & Kien, 1997). These studies measured the average duration of sub-behaviors within the activities to establish the rate at which mammals segment motor behaviors. Although the findings concurred on a fixed, cross-modal and cross-cultural character of the segmentation, the actual length has remained controversial. Whereas Pöppel built his theory around what he called a three-second window that he argued applies to all sorts of behaviors (2004), Schleidt and Kien (1997) suggested a 1-4-second unit and showed that the length differed across behaviors.

These estimations pertain mostly to behavior segmentation in humans, the most extensively studied mammal to date. As such, humans are a good reference species for animals whose behavioral segmentation has never been studied before. Therefore, when investigating behavior segmentation in an avian species, the common raven (*Corvus corax*), for the first time, we compared it directly to humans in a similar set up. We also extended previous research by including a choice task: making a simple choice between non-functional objects. We hypothesize that the length of the behavior segments, in particular sensorial ones, may influence the speed of basic cognition in certain tasks, which in turn could affect the speed of solving more complex tasks.

In species with vision as their primary sensory modality, such as ravens and humans, eye gazes are arguably indicative of information gathering processes (Davidson, Butler, Fernández-Juricic, Thornton, & Clayton, 2014), and gaze behaviors may be a good source of data when comparing behavioral segmentation in two such phylogenetically distant species. Therefore, to estimate a difference in behavioral segmentation, we compared the length of the single gazes (hereafter, gaze durations for brevity) in humans and ravens in the same task.

To establish reliable criteria for comparison, we considered differences in the visual configuration of humans and ravens, as they may contribute to differences in gaze duration between these two groups. Despite some similarities in eye anatomy, ravens differ from humans with respect to visual field and flicker-fusion frequency, and these differences are crucial for our predictions. Ravens, just like humans, have one fovea (a region of acute vision) in their retina (Martin, 2017; Wood, 1917). It is responsible for acute vision within the binocular overlap, which is smaller in ravens (43.2 ± 2.4, Troscianko, Bayern, Chappell, Rutz, & Martin, 2012) than in humans (114 degrees). Humans' frontal location of eyes in the skull – as opposed to lateral in ravens – trims the rear parts of lateral visual field, restricting it to only its frontal area. Although these parameters underpin differences in ravens' and humans' orientation when attending to gaze cues, we expect that, in our task with static and proximate visual stimuli, both the ravens and the humans will predominantly use frontal visual field (over lateral visual field in the ravens; Maldonado, Maturana, & Varela, 1988). Moreover, while gazing at a stimulus, ravens will most likely lock their head in one position (Maldonado et al., 1988) and humans may both lock head and change its position in a continuous manner.

Humans and ravens most likely differ extensively in their vision's temporal resolution, i.e., in how much light information they can perceive in a given time unit, which is estimated by critical flicker-fusion frequency (CFF): the highest flicker-fusion-frequency at any light intensity (Healy, McNally, Ruxton, Cooper, & Jackson, 2013; Lisney et al., 2011). Simply put, flicker-fusion-frequency is the frequency with which a flickering light is perceived as a continuous steady light flow. This is measured in numbers of flickers per second and described in Hz. Although the precise rate for ravens is yet to be established, based on studies with other songbird species (blue tit, collared flycatcher, pied flycatcher: Boström, Dimitrova, Canton, Håstad, Qvarnström, & Ödeen, 2016; European starlings: Greenwood et al., 2004), it is likely that ravens' CFF exceeds 100 Hz. Humans on the other hand average around 60 Hz. In other words, when a series of light flashes is displayed to a human, s/he would notice the delays between consecutive flashes only when fewer than 60 per second, while a songbird would need more than 100 flickers per second before the light is perceived to stop blinking. The difference in CFF results from different adaptations in humans and ravens related to their locomotion (Jones, Pierce, & Ward, 2007). Flight requires quick, agile responses to the changes in the optic flow, which is usually much slower in humans' terrestrial environment.

It is of interest to comparatively study a mammal and a bird species in a simple task with a focus on their visual behavior. Little is known, for instance, whether and how differences in perceptual abilities

constrain birds' and mammals' performance in physically similar setups. For example, comparative memory studies often introduce a delay between encoding and retrieval of the target information, but with the current state of knowledge, it is difficult to assess whether an objectively identical delay is comparable between birds' and mammals' memory storage skills. If the time flies faster for birds, more time would subjectively pass than it would for mammals during an identical delay. Therefore, the birds would, in a sense, have to store the target information for a longer time before retrieval than mammals, and, during this delay, they may hence be more susceptible to various memory distortions. Here we make an initial study to approach the issues of potential differences in perceptual speed when engaged in a simple choice task.

We used a forced-selection task to investigate differences in gaze duration between humans and ravens. The task was chosen because the ravens had extensive experience of making forced selections of objects from a board (i.e., that only one out of several may be selected; e.g., Kabadayi & Osvath, 2017; and in several unpublished studies). This made it likely that they would pay attention to the objects before they made a selection. Another benefit of this setup is that ravens very rarely refuse to select non-functional objects, as they have a highly pronounced interest in interacting with them (Jacobs et al., 2014). The humans were explicitly informed that they must choose only one of the target objects. We therefore posited that, while engaged in the task, the subjects made a simple decision before selecting an object. If our definition of a behavior segment was in line with those used before in humans and primates, we expected that the gaze durations should have a right-skewed distribution; that is, fewer longer than shorter gazes would be detected in both species (Schleidt & Kien, 1997). Furthermore, based on the differences in visual systems, we expected that the median gaze duration in humans would be significantly higher than in ravens, and that would reach between 1 and 4 seconds, based on Schleidt's and Kien's report (1997). This relatively shorter duration for humans, compared to the other studies, is more likely, as Schleidt and Kien (1997) measured behavioral segmentation in human preparatory/planning behaviors preceding interactions with objects. Specifically, when subjects were asked to group the objects according to certain physical features, their motor actions typically lasted for 2.4-2.5 seconds. Although this study was similar to ours in the sense that it involved exposure to objects, it focused on motor and not perceptual aspects.

Method

Subjects

Twenty adult humans (11 females) of multiple ethnical backgrounds (M = 33 years, age range: 22 – 52 years) and five adult ravens (2 males) participated in the study. The humans were tested individually at Lund University; they were students or employees of the university and of international background, and from different academic disciplines. The humans were familiar with the experimenter, and were not rewarded for their participation; the motivation provided was to contribute to science. The ravens were housed at the Lund University Corvid Cognition Station in a social group in a 400 m^2 space. The tests were conducted individually in familiar facilities with free access to food and water during the experiment. The ravens were familiar with the experimenter.

Stimuli

Gazes at familiar objects have been shown to be shorter than at non-familiar ones (e.g., in human infants, Houston-Price & Nakai, 2004). A control for this effect was necessary, as stimuli presented in this study could vary in degree of familiarity to subjects. Therefore, two sets of objects were used in the study: one with familiar and easily categorized objects (e.g., a spoon for humans and a bottle cap for ravens), and another with novel objects that arguably are difficult to immediately categorize (objects that were manufactured only for this study). In half of trials, only familiar or only non-familiar objects were

presented simultaneously, while in the other half, both familiar and non-familiar objects were presented (for details see Procedure).

Because humans and ravens interact with different objects on a daily basis, they were tested with different sets (Figures 1-2). The objects were sized for each species such that they could be held in the beak or hands of the subjects. Each of the four sets contained 20 different objects, out of that, 18 were randomly chosen for each subject. The level of complexity of visual characteristics of the familiar and the non-familiar objects was similar within species. As the familiar objects for the ravens had simpler shapes and were usually monochromatic, the non-familiar objects followed the same schema. As the familiar objects for the humans were already more complex than the familiar objects for the ravens, i.e., had more intricate shape details and often more than one color, the non-familiar objects for the humans were also more visually complex than those for the ravens.

Figure 1. A display of sample objects from the sets for the ravens: A) familiar, B) non-familiar.

Figure 2. A display of sample objects from the sets for the humans: A) familiar, B) non-familiar.

Apparatus

The apparatus consisted of a low table (120 x 30 x 23.5 cm) with a flat tray (100.5 x 30 x 1.3 cm). Hidden underneath the table between the trials, the tray served the presentation of stimuli during the trial. Each subject was tested in a selection procedure (humans: Figure 3, ravens: Figure 4). The objects were equidistantly placed on the tray (humans: 30 cm between objects, ravens: 10 cm between objects). This set of objects consisted either of only familiar or non-familiar items, or a combination of the two categories.

Figure 3. A display of testing procedure for the humans. A) Before each testing session, twelve sets of three objects were hidden underneath a low table. B) During each trial, one set of objects was presented on a tray to a subject.
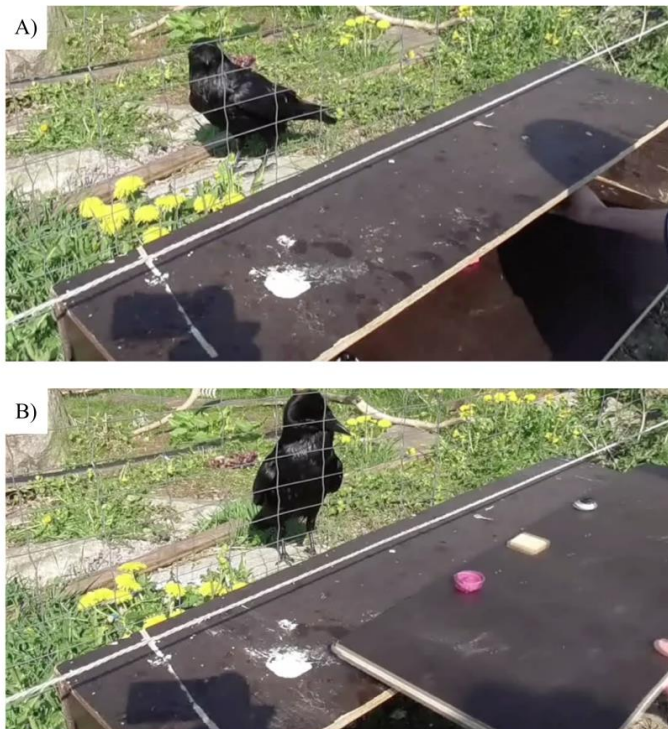

Figure 4. A display of testing procedure for the ravens. A) Before each testing session, twelve sets of three objects were hidden underneath a low table. B) During each trial, one set of objects was presented on a tray to a subject.

Procedure

The humans were individually tested in a single session of 12 trials, and the ravens were individually tested in 1-3 sessions, also adding up to 12 trials. Four combinations of the objects were possible (3x familiar, 3x non-familiar, 2x familiar and 1x non-familiar, 1x familiar and 2x non-familiar), and each combination was repeated three times within the 12 trials. The order of trials was pseudorandomized. For the humans, the selection procedure was preceded by an instruction (Appendix 1)

and for the ravens, it was preceded by a familiarization to the apparatus. The ravens were already familiar with the forced choice paradigm; that is, that only one object may be selected per trial. The presentation of the selection tray functioned as an offer to select, and they never refused to do so. Obviously, this differs from a linguistic instruction, but based on our previous experience that they always select (and only one item), even when the objects are not rewarded and even consist of everyday objects part of their aviary (e.g., Kabadayi & Osvath, 2017), we assumed that this was as close as we could get to an instruction. To ensure that the subjects' eyes are visible throughout the testing, the humans sat on a chair behind the table and the ravens were separated from the apparatus by a mesh, whose openings were wide enough to allow for the object retrieval.

At the beginning of the trial, all objects were hidden from the subject. First, the experimenter placed the objects equidistantly on the tray out of sight of the subject, then placed the tray on the table in view of the subject, and pushed the tray within the subject's reach. The subject could then select an object by grasping it with hand or beak. None of the choices were reinforced. After the subject selected an object, the tray was removed, hidden and emptied. This procedure was repeated each trial. Each session was video recorded.

Neither the humans nor the ravens were extrinsically rewarded (monetary compensation or food) for participating in the study. Based on previous research on the same birds (Jacobs et al., 2014), we expected that interacting with objects would be rewarding in itself for the ravens. The humans were always tested by a familiar experimenter and explicitly informed in the instruction that there were neither right nor wrong choices.

Coding

Davidson and others (2014) reported that bird species, which have laterally placed eyes and a single fovea like ravens, move their heads more than their eyes when changing the direction of gaze. This is why gaze fixations in birds have been also described as "gaze locking" (Butler, Templeton, & Fernandez-Juricic, 2018). Within this behavior, a bird locks the visual input for a certain period of time with a lateral view (monocular gaze locking) or frontal view of the object (binocular gaze locking). Based on this information, combined with Maldonado and colleagues' report (1988), we expect that ravens should gaze at static and proximate objects frontally, without changing head position.

The gazes on the objects were video analyzed. The criteria for defining a single gaze at an object were:

1. The gaze of at least one eye (ravens) or both eyes (humans) were directed towards an object.
    a. Ravens' beak was directed toward the object; head position was not changing throughout the gaze.
    b. Humans' head position did not matter as long as eyes were directed towards the object.
2. The whole length of the gaze was observable (from onset to offset).
    a. In ravens, the gaze started once head was locked and all criteria from 1a were fulfilled. The gaze ended when head's position changed.
    b. In humans, the gaze started once eyes were fixated on the object and ended when eyes moved away from the object.
3. Gazes were coded only when the subject was not locomoting.

The durations of gazes were coded frame-by-frame in ELAN 4.9.3. The coding started once the tray was leveled with the top surface of the table, allowing the subject to look at all three objects. The coding ceased with the subject's last gaze before selecting an object. The recordings were pooled for coding by two independent raters. Inter-observer agreement was high (0.865 in ravens, 0.828 in humans; time-unit kappa without tolerance for 23.6% of trials, Appendix 1). A margin of mistake in coding was set to 0.2 s.

Statistics

Neither the raven nor the human data fit a normal distribution (Shapiro-Wilk test, $p < .001$). The best fitting distributions were determined with qqp function from MASS package in R. A general linear mixed-model (GLMM) analysis was used to determine the effect of several explanatory variables (species, trial number, stimulus familiarity) on a response variable (gaze duration; single instances or total gaze duration per trial, or average gaze duration per trial), controlling for an effect of a random variable (subject ID). Functions glmer and glmmPQL were used from packages lme4 and mlmRev, respectively. All statistical analyses were conducted in *R* (v.3.5.1, the R Foundation for Statistical Computing: http://www.R-project.org). Significance level was set at .05 and Bonferroni correction was used for all p-values for significant effects within the models.

Results

Gaze Durations

Log-normal distribution best fitted the gaze durations in both the ravens and the humans. GLMM analysis with the gaze durations with a response variable and three explanatory variables – species, object familiarity and trial number, with subject ID as a random variable – revealed significant main effects of species ($\chi^2(1) = 54.95$, $p < .001$) and object familiarity ($\chi^2(1) = 22.49$, $p < .001$) on the gaze durations, and a significant interaction effect ($\chi^2(3) = 80.59$, $p < .001$). Overall, the ravens cast shorter gazes at the familiar ($z = 5.24$, $p < .001$) and the non-familiar ($z = 6.66$, $p < .001$) objects than the humans. Furthermore, in the ravens there was no difference in the gaze durations for familiar and non-familiar objects ($z = 0.02$, $p = 1$), but there was a significant difference in the humans ($z = -4.85$, $p < .001$; Figure 5).
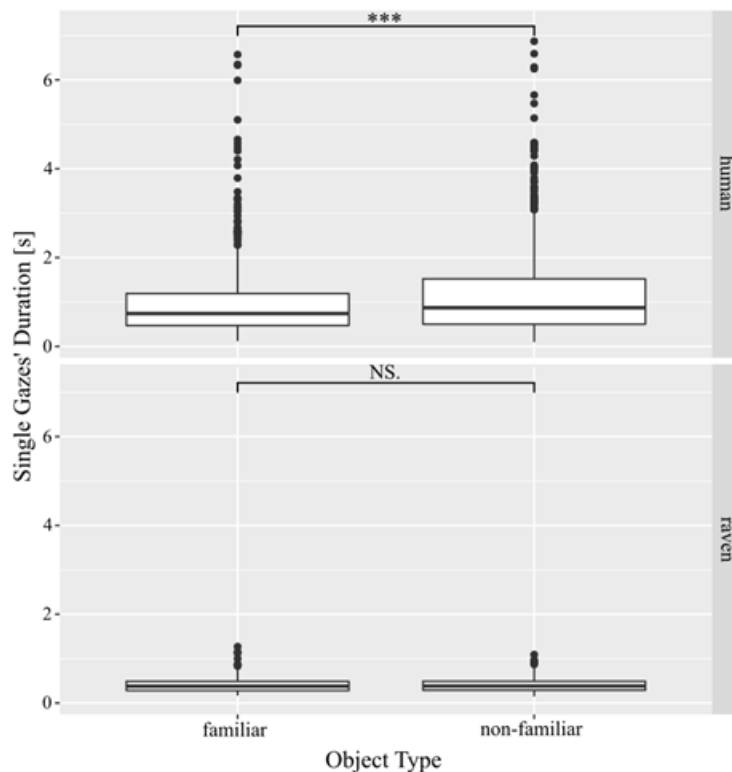

Figure 5. The effect of species and object type on gaze durations.

Therefore, the ravens' and the humans' the gaze durations were calculated separately for the familiar objects (ravens: *Mdn* = 0.38, median absolute deviation *(MAD)* = 0.148, interquartile range *(IQR)* = 0.282-0.488; humans: *Mdn* = 0.74, *MAD* = 0.474, *IQR* = 0.47-1.19) and for the non-familiar objects (ravens: *Mdn* = 0.384, *MAD* = 0.142, *IQR* = 0.29-0.49; humans: *Mdn* = 0.87, *MAD* = 0.652, *IQR* = 0.5-1.52; Figure 6).
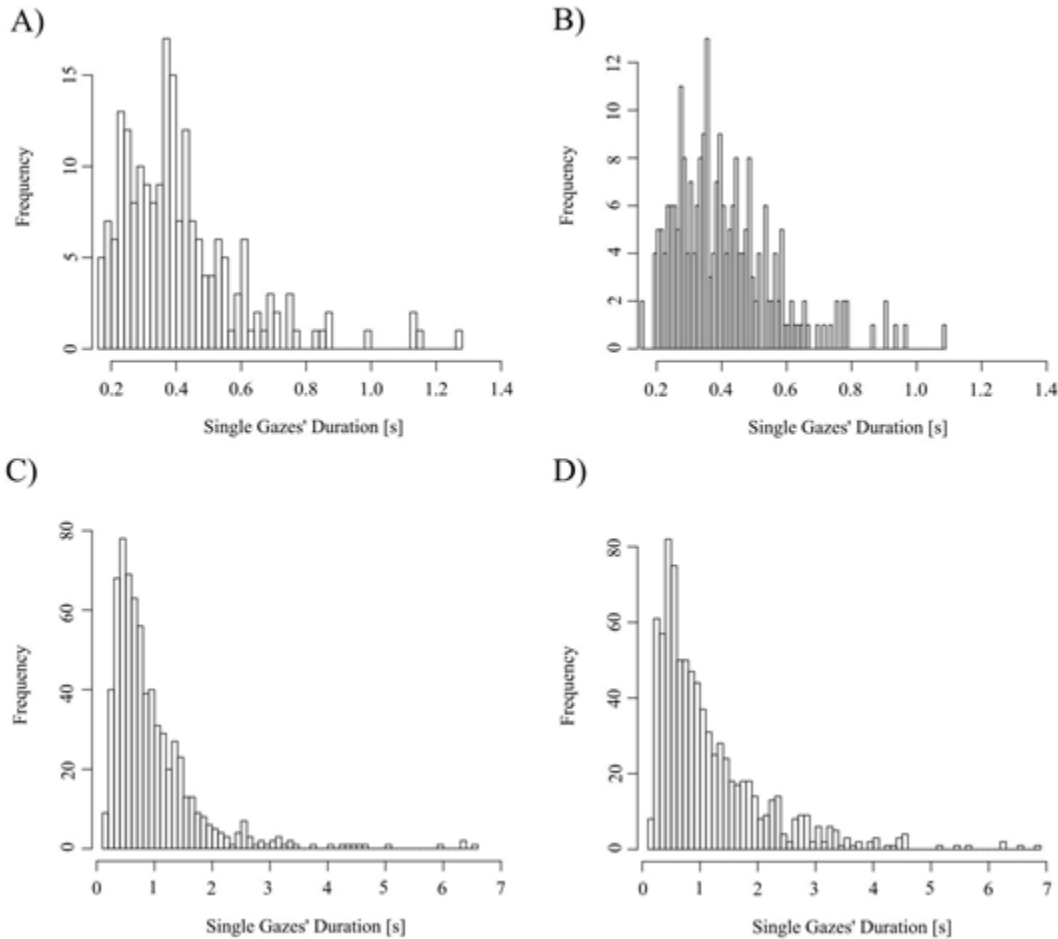

Figure 6. A histogram of gaze durations for each species and each object type: A) for the ravens and the familiar objects, B) for the ravens and the nonfamiliar objects, C) for the humans and the familiar objects, and D) for the humans and the nonfamiliar objects.

Number of Single Gazes

Log-normal distribution best fitted the number of gaze durations per trial in both the ravens and the humans.

Based on GLMM analysis with the number of gaze durations per trial as a response variable, species and trial number as explanatory variables, and subject ID as a random variable, a main effect of species was not significant ($\chi^2(1)$ = 0.468, *p* = .494). However, both a main effect of trial ($\chi^2(11)$ = 126.949, *p* < .001; Figure 7) and an interaction effect for species and trial ($\chi^2(11)$ = 43.709, *p* < .001) were significant. Overall, the individuals of both species cast the lowest number of gazes before grabbing one of the objects in the 8th, 11th and in the 12th trial and the highest number of gazes in the 4th trial. The number of gazes in the 8th trial was significantly lower than in the 1st (z = -355, *p* = .018), the 2nd (z = -361, *p* = .015), the 4th (z = -5.15, *p* < .01), the 5th (z = -4.07, *p* < .01) and the 7th trial (z = -3.68, *p* = .012);

it was also significantly higher in the 4th trial than in the 3rd ($z = 3.96$, $p < .01$), the 6th ($z = -4.6$, $p < .01$), the 9th ($z = -3.73$, $p < .01$), the 10th ($z = -4.26$, $p < .01$), the 11th ($z = -4.69$, $p < .01$) and the 12th trial ($z = -4.88$, $p < .01$). Some significant differences were also detected between trials for each species separately. In the ravens, there was a significantly higher number of gazes in the 1st trial than in the 2nd ($z = -4.16$, $p < .01$), the 3rd ($z = -3.54$, $p = .46$), the 5th ($z = -4.2$, $p < .01$), the 6th ($z = -4.27$, $p < .01$), the 7th ($z = -4.21$, $p < .01$), the 8th ($z = -4.08$, $p < .01$), the 9th ($z = -3.85$, $p = .016$), the 10th ($z = -3.93$, $p = .011$) and the 11th trial ($z = -4.12$, $p < .01$). In the humans, there was a significantly higher number of gazes in the 4th trial than in the 3rd ($z = 3.96$, $p = 0.011$), the 6th ($z = -4.6$, $p < .01$), the 8th ($z = -5.15$, $p < .01$), the 9th ($z = -3.72$, $p = .025$), the 10th ($z = -4.26$, $p < .01$), the 11th ($z = -4.69$, $p < .01$) and the 12th trial ($z = -4.88$, $p < .01$).
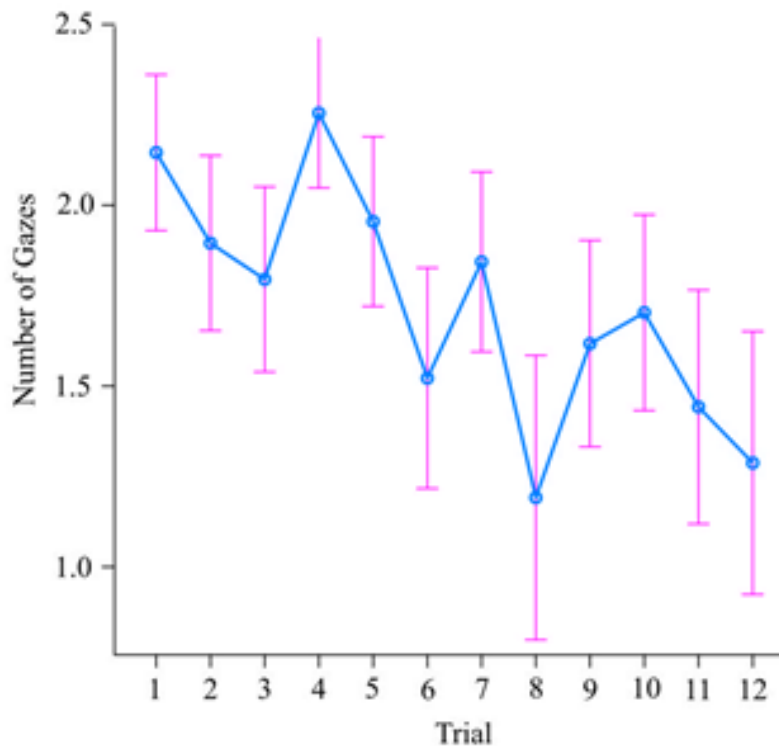

Figure 7. An effect of trial on the number of gazes per trial. As there was no main effect of species, this is a display for both the ravens' and the humans' gazes combined.

Gamma distribution best fitted the number of gaze durations per object type (familiar vs. non-familiar) both in the ravens and in the humans.

Based on GLMM analysis with the number of gaze durations per object type as a response variable, species and object type as explanatory variables, and subject ID as a random variable, a main effect of object type was significant ($\chi^2(1) = 7.14$, $p = .008$; Figure 8). A main effect of species ($\chi^2(1) = 0.62$, $p = .432$) and an interaction effect of species and object type ($\chi^2(1) = 0.56$, $p = .455$) were not significant. Specifically, the individuals of both species cast higher number of gazes at non-familiar than for the familiar objects ($z = 2.72$, $p = .006$). For the familiar objects, the ravens cast 40.4 gazes on average ($SD = 14.05$) and the humans cast 35 gazes on average ($SD = 17.13$). For the non-familiar objects, the ravens cast 48 gazes on average ($SD = 34.36$) and the humans cast 41.45 gazes on average ($SD = 25.48$).
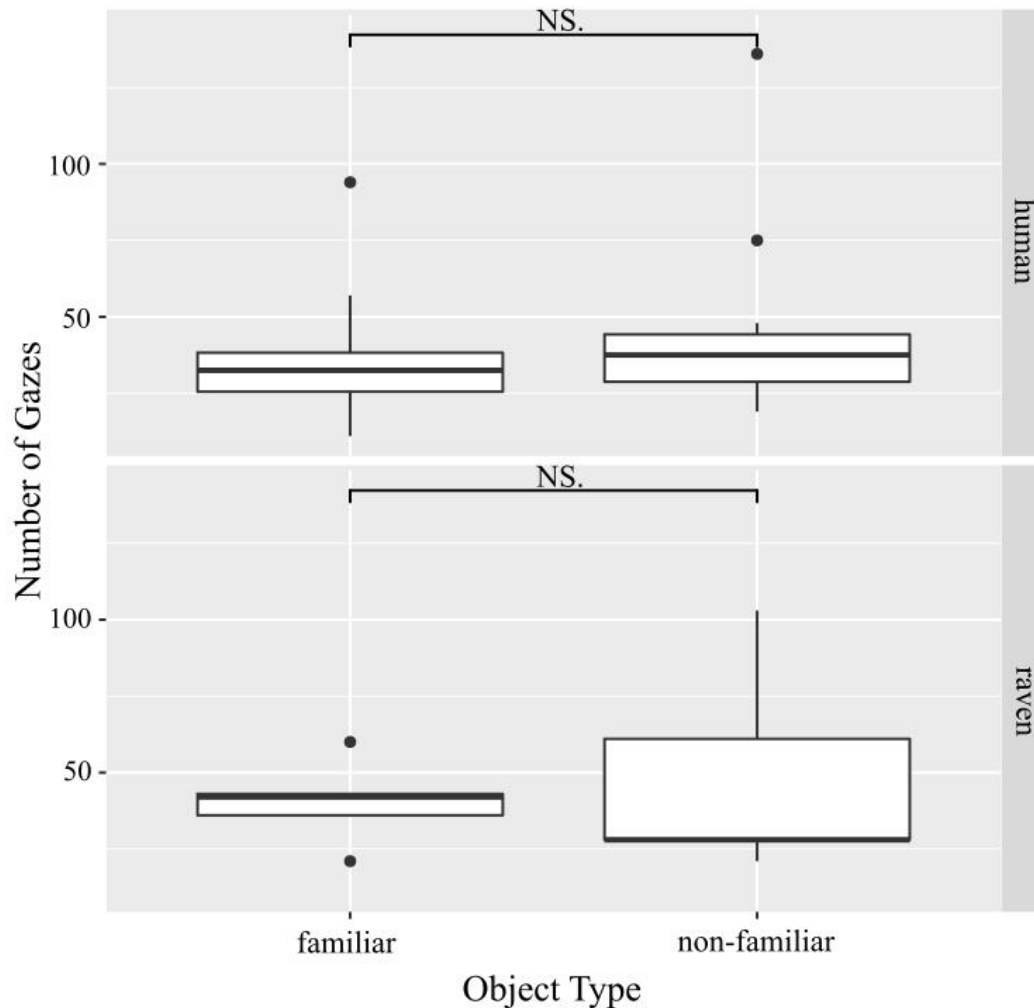
Figure 8. An effect of object type and species on the number of gazes per object type.

Total Gaze Duration

Log-normal distribution best fitted total gaze duration in both the ravens and the humans.

Based on GLMM analysis with total gaze duration as a response variable, species and trial number as explanatory variables, and subject ID as a random variable, a significant main effect of trial number ($\chi^2(11) = 136.317$, $p < 0.001$) on total gaze duration per trial was found, but the main effect of species ($\chi^2(1) = 4.08$, $p = 0.08$) and the interaction effect for these variables ($\chi^2(11) = 11.93$, $p = 0.369$) were not significant. Specifically, there was a statistically significant effect of trial on total gaze duration in both the ravens ($\chi^2(11) = 34.54$, $p < 0.001$; Figure 9Aa), and the humans ($\chi^2(11) = 117.27$, $p < 0.001$; Figure 9Bb).
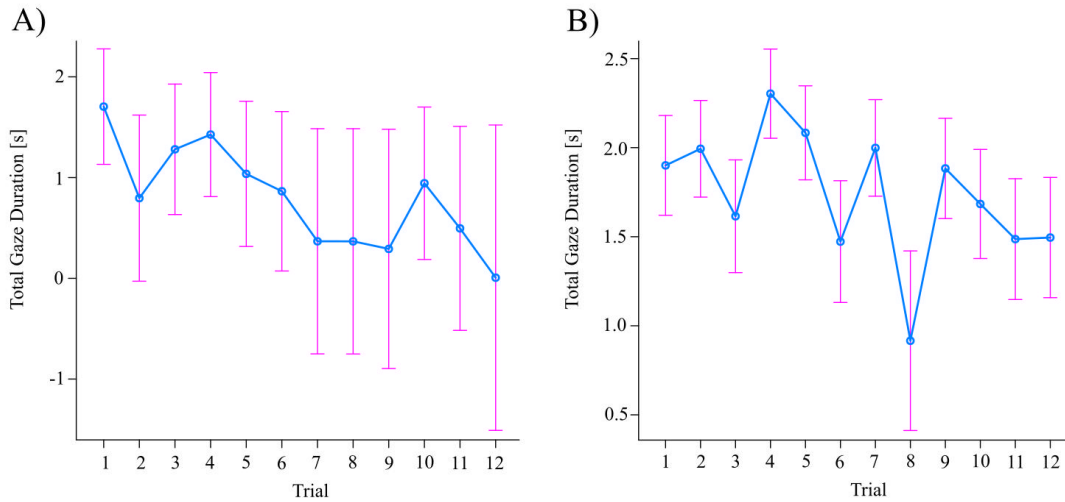
Figure 9. An effect of trial on total gaze duration: A) in the ravens, B) in the humans.

Average Gaze Duration per Trial

Gamma distribution best fitted average gaze duration per trial in both the ravens and the humans. Based on GLMM analysis with average gaze duration as a response variable, species and trial number as explanatory variables, and subject ID as a random variable, only a main effect of species was significant ($\chi^2(1) = 122.991$, $p < .001$), contrary to a main effect of trial ($\chi^2(11) = 14.45$, $p = .21$); no interaction effect was detected ($\chi^2(11) = 8.71$, $p = .65$). The average gaze duration in the humans was significantly higher than in the ravens ($z = -11.03$, $p < .001$), with a mean of 1.09 s for the humans ($SD = 0.478$) and a mean of 0.411 for the ravens ($SD = 0.094$; Figure 10).
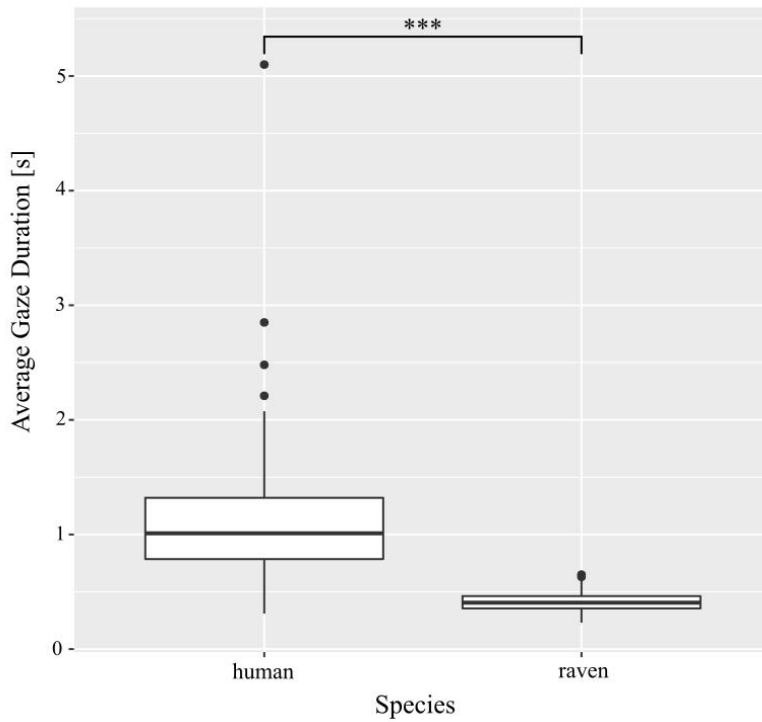

Figure 10. An effect of species on average gaze duration per trial.

Cumulative Gaze Duration

Cumulative distributions of gaze durations for the ravens and for the humans are presented in Figures 11 and 12. Overall, the distributions both for the ravens and for the humans had a similar shape: quite flat, with a steady increase across the trials.
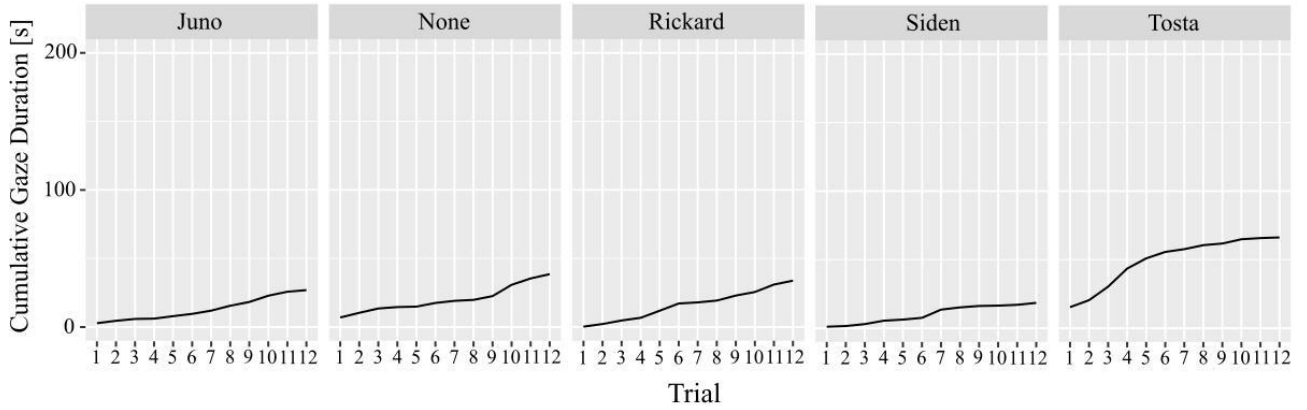

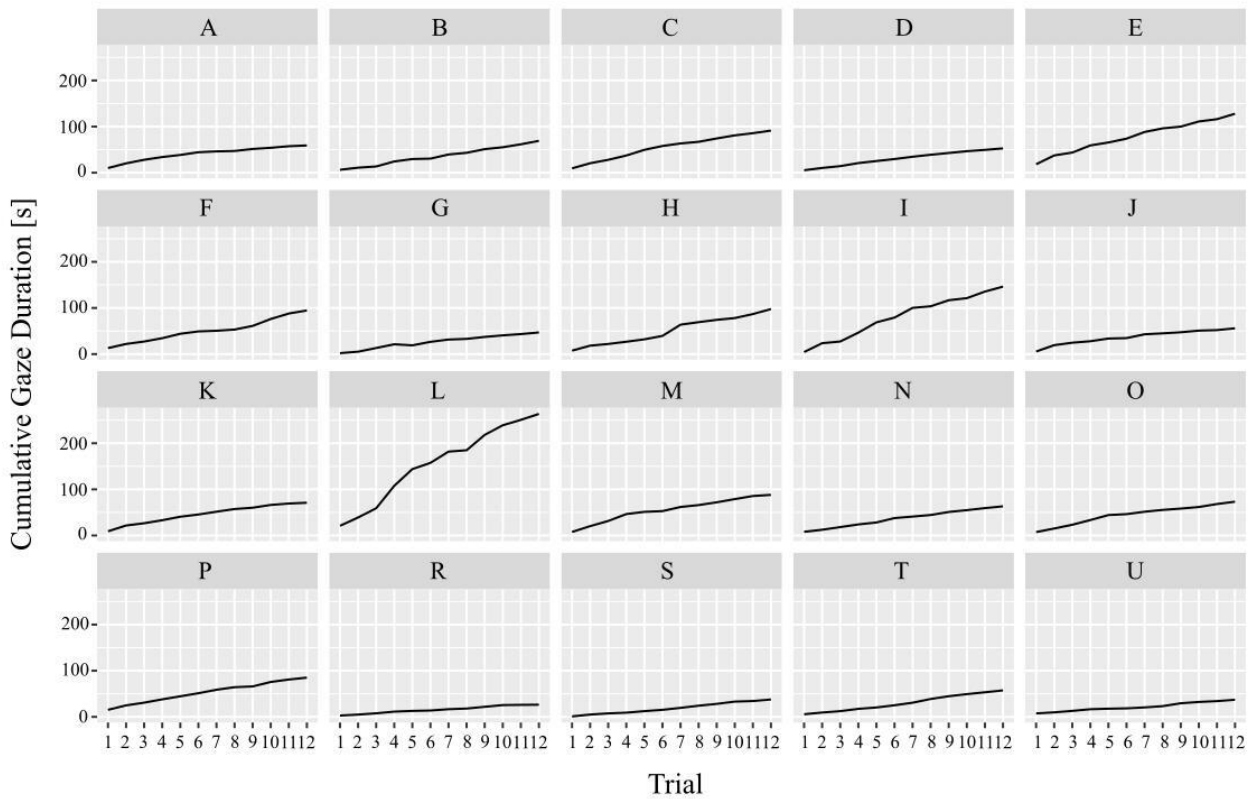*Figure 11.* Cumulative distributions of gaze duration in the ravens.

Figure 12. Cumulative distributions of gaze duration in the humans.

Discussion

In line with our expectations, the ravens gazed at both the familiar and the non-familiar objects for approximately half of the duration compared to the humans, and the durations followed the right-skewed distribution in both groups. However, we detected a shorter unit in the humans than reported in previous studies (0.74 and 0.87 s vs. 1-4 s). Those studies predominantly investigated motor rather than perceptual patterns, which may differ in timing. Motor patterns have perhaps a slower timing due to neurological processes that ensue perception and precede action, or due to engaging more muscle units and larger body parts as compared to eye movements.

Furthermore, we detected a difference in the gaze durations between the familiar and the non-familiar objects in the humans, but no such difference in the ravens. The human results are consistent with previous reports of the effect of novelty on visual perception (Houston-Price & Nakai, 2004). The perception of the non-familiar objects likely required additional cognitive processing which increased gaze duration in the humans. We did not detect this effect in the ravens, which may have resulted from the lack of the habituation to any of the objects before testing, and/or the lack of cost associated with the choice – the task was a rather simple one. Therefore, any potential categorization time in the ravens could not be detected. As we detected a novelty effect only in the humans, we compared the gaze durations between the species for each object type separately.

The observed differences could be explained in at least four ways: (1) the visual characteristics of the objects presented to the humans were more complex than of those presented to the ravens; (2) the humans could have been more eager to perform "well" in the task and paid more attention to the presented objects; (3) both species suffered from a decrease in attention over trials, but the humans suffered from a smaller decrease in attention than the ravens; decrease in attention could lead to shorter gaze durations, and if such a decrease was larger in the ravens, their gazes should be significantly shorter than the humans; (4) the humans needed more time to collect relevant visual information than the ravens before selecting an object.

If the first explanation is correct, one would expect that the humans should have cast (1a) longer single gazes than the ravens, and (1b) had a higher number of single gazes than the ravens before making a choice, and (1c) that a difference in gaze duration for the object types, or a lack of such difference, should be the same in both species. More visually complex objects would require acquiring more visual information and more processing of this information before making the decision; therefore, one would expect that the humans would maximize their intake of visual information by longer gazes and a higher number of gazes as opposed to the ravens, exposed to less visually complex objects. Besides, if one assumes that both familiar and nonfamiliar objects were similar in the level of visual complexity for each species, and that visual complexity had an effect on gaze duration, one would also expect in both species either a difference in gaze duration between the object types in both species or a lack of such difference. Only the first expectation (1a) was consistent with our results.

If the second explanation is correct, one would have expected that the humans should have cast (2a) longer single gazes than the ravens, (2b) and more gazes than the ravens before making the selection. If one were more motivated to perform "well", one would be eager to acquire more visual information than if one was less motivated, in order to make the best possible choice; acquiring of additional visual information would in turn require longer and more gazes from the humans than from the ravens before making the selection. Although the first expectation (2a) was consistent with our results, the ravens in fact (2b) cast more gazes on the objects regardless of the object type than the humans.

If the third explanation is correct, one would expect that (3a) the number of gazes per trial should have decreased over trials, but significantly more in the humans than in the ravens, (3b) total gaze duration per trial should have decreased over consecutive trials, but significantly more in the humans than in the ravens, and (3c) average gaze duration per trial should have also decreased over consecutive trials, but significantly more in the humans than in the ravens. If one was forced to select a visually distinctive item at as low attention costs as possible, one should minimize effort of acquiring visual information. This means that loss in attention should result in reducing the duration and the number of single gazes,

and as a result, also the total gaze duration per trial. When the task was still new, that is in the first few trials, one could assume that the subjects would pay more attention to the objects than in the following trials, when the novelty of the task would have faded away. All subjects, regardless of the species, could have suffered from this effect; however, the ravens' gazes tended to be shorter than the humans' and so one could assume that the decrease in attention was higher in the ravens than in the humans. Whereas the first two expectations (3a and 3b) were supported by our results, the third one (3c) was not: in fact, there was no effect of trial number on average gaze duration per trial in the two tested species. Therefore, it seems that the subjects, regardless of the species, attempted to speed up the object selection, but did so by reducing the number of gazes, not their duration. They might not have been able to control, and therefore reduce, their gaze duration.

The fourth explanation accommodates all of the above findings, as it states that the humans needed more time to collect relevant visual information than the ravens before selecting an object. We suggest that this difference has neurobiological and sensory system causes. Ravens have a higher neuronal density of pallial neurons than humans (Olkowicz et al., 2016), which might increase processing speed in simple perceptually based decisions. Like other passerines, ravens also most likely have a notably higher temporal resolution in their vision compared to humans, which probably reduces the time to collect relevant information before a decision is made. This is highly adaptive for a flying diurnal species, as the high temporal resolution allows for rapid perception of fast-moving objects and quick responses to ever-changing optic flow. In slow-paced conditions, such adaptations may be observed as more quickly executed behaviors, compared to a terrestrial species.

The speed of temporal perception most likely has a matching speed of cognitive processing (Shaw & Bransford, 1977). One could speculate that if such speed is combined with a complex brain with high neuronal numbers, it might give an additional cognitive effect not dictated by the numbers alone. Ravens are often found to parallel apes in their cognitive abilities (Güntürkün & Bugnyar, 2016). Although many cognitive studies are interested in a binary outcome rather than temporal aspects of the performance, such aspects could provide additional information about how a task is solved, and how proficiently, and allow for comparisons between species. Temporal aspects of performance may contribute to a more exhaustive analysis of a variety of cognitive tasks, such as detour tasks, causal cognition tasks and other.

The results may have implications for comparative memory research, as it often involves a delay between the encoding and the retrieval of stimuli. As attentional mechanisms enforce different frame rates on the perception of the actual stimuli between the species, they might also, indirectly, affect performance of the tested species after a fixed delay. Although such delay between the initial exposure and testing may be identical for the external observer, it is non-identical from the perspective of the tested individual. Animals that process more information per a fixed time interval might be more cognitively taxed than animals processing less because, arguably, more has to be encoded in their memories; and in a sense, the subjective time that has passed becomes longer. That is, an identical absolute time interval duration may be relatively longer for animals that process more information than for those that process less. Moreover, it has previously been shown that birds, like mammals (rats: Kraemer, Brown, & Randall, 1995; humans; Sucala, Scheckner, & David, 2010), make decisions based on relative rather than absolute interval durations (pigeons: Zentall, Weaver, & Clement, 2004). As a result, the same delay between exposure and testing might, in effect, be experienced differently in birds and mammals, which in comparative studies between these groups calls for controls of such factors.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. The facility and the care taking routines were approved by the Swedish Agricultural Board (No 5.2.18-5395/16). Ethical

approval for the procedures was granted by the regional ethics board for animal research in the county of Skåne (No M 333-12).

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

Both authors contributed equally and agreed to be held accountable for the contents of this work and approved the final version of the manuscript.

Acknowledgments

References

Baer, K. E. R. v. (1862). *Welche Auffassung der lebenden Natur ist die richtige? Festrede zur Eröffnung der russischen entomologischen Gesellschaft im Mai 1860*. Berlin.

Boström, J. E., Dimitrova, M., Canton, C., Håstad, O., Qvarnström, A., & Ödeen, A. (2016). Ultra-rapid vision in birds. *PLoS ONE, 11*, e0151099. http://dx.doi.org/10.1371/journal.pone.0151099

Butler, S. R., Templeton, J. J, & Fernandez-Juricic, E. (2018). How do birds look at their world? A novel avian visual fixation strategy. *Behavioral Ecology and Sociobiology, 72*, 38. http://dx.doi.org/10.1007/s00265-018-2455-0

Davidson, G. L., Butler, S., Fernández-Juricic, E., Thornton, A., & Clayton, N. S. (2014). Gaze sensitivity: Function and mechanisms from sensory and cognitive perspectives. *Animal Behaviour, 87*, 3–15. http://dx.doi.org/10.1016/j.anbehav.2013.10.024

Gerstner, G. E., & Goldberg, L. J. (1994). Evidence of a time constant associated with movement patterns in six mammalian species. *Ethology and Sociobiology, 15*, 181–205. http://dx.doi.org/10.1016/0162-3095(94)90013-2

Gibson, J. J. (1975). Events are perceivable but time is not. In J. T. Fraser & N. Lawrence (Eds.), *The study of time II* (pp. 295–301). New York, Heidelberg, Berlin: Springer Verlag.

Greenwood, V. J., Smith, E. L., Goldsmith, A. R., Cuthill, I. C., Crisp, L. H., Walter-Swan, M. B., & Bennett A. T. D. (2004). Does the flicker frequency of fluorescent lighting affect the welfare of captive European starlings? *Applied Animal Behaviour Science, 86*, 145–59. http://dx.doi.org/10.1016/j.applanim.2003.11.008

Güntürkün, O., & Bugnyar, T. (2016). Cognition without cortex. *Trends in Cognitive Sciences, 20*, 291-303. http://dx.doi.org/10.1016/j.tics.2016.02.001

Healy, K., McNally, L., Ruxton, G. D., Cooper, N., & Jackson, A. L. (2013). Metabolic rate and body size are linked with perception of temporal information. *Animal Behaviour, 86*, 865–896. http://dx.doi.org/10.1016/j.anbehav.2013.06.018

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development, 13*, 341–348. http://dx.doi.org/10.1002/icd.364

Jacobs, I. F., Osvath, M., Osvath, H., Mioduszewska, B., Bayern, A. M. P. v., & Kacelnik, A. (2014). Object caching in corvids: Incidence and significance. *Behavioural Processes, 102*, 25–32. http://dx.doi.org/10.1016/j.beproc.2013.12.003

Jones, M. P., Pierce K. E., & Ward, D. (2007). Avian vision: A review of form and function with special consideration to birds of prey. *Journal of Exotic Pet Medicine, 16*, 69–97. http://dx.doi.org/10.1053/j.jepm.2007.03.012

Kabadayi, C., & Osvath, M. (2017). Ravens parallel great apes in flexible planning for tool-use and bartering. *Science, 357*, 202–204. http://dx.doi.org/10.1126/science.aam8138

Kraemer, P. J., Brown, R. W., & Randall C. K. (1995). Signal intensity and duration estimation in rats. *Behavioural Processes, 34*, 265–268. http://dx.doi.org/10.1016/0376-6357(95)00003-D

Lisney, T. J., Rubene, D., Rózsa, J., Løvlie, H., Håstad, O., & Ödeen, A. (2011). Behavioural assessment of flicker fusion frequency in chicken *Gallus gallus domesticus*. *Vision Research, 51*, 1324–1232. http://dx.doi.org/10.1016/j.visres.2011.04.009

Maldonado, P. E., Maturana, H., & Varela, F. J. (1988). Frontal and lateral visual system in birds. Frontal and lateral gaze. *Brain, Behavior and Evolution, 32*, 57–62. http://dx.doi.org/10.1159/000116532

Martin, G. R. (2017). *The sensory ecology of birds*. Oxford, UK: Oxford University Press.

Olkowicz, S., Kocourek, M., Lucan, R. K., Porteš, M., Fitch, W. T., Herculano-Houzel, S., & Němec, P. (2016). Birds have primate-like numbers of neurons in the forebrain. *Proceedings of the National Academy of Sciences of the United States of America, 113*, 7255-7260. http://dx.doi.org/10.1073/pnas. 1517131113

Pöppel, E. (2004). Lost in time: A historical frame, elementary processing units and the 3-second window. *Acta Neurobiologiae Experimentalis, 64*, 295–301.

Schleidt, M., & Kien, J. (1997). Segmentation in behaviour and what it can tell us about brain function. *Human Nature, 8*, 77–111. http://dx.doi.org/10.1007/s12110-997-1005-7

Shaw, R., & Bransford, J. (1977), *Perceiving, acting and knowing. Towards an ecological psychology*. Hillsdale, NJ: Lawrence Erlbaum.

Sucala, M., Scheckner, B., & David, D. (2010). Psychological time: Interval length judgments and subjective passage of time judgments. *Current Psychology Letters, 26*, 1–9.

Troscianko, J., Bayern, A. M. P. v., Chappell, J., Rutz, C., & Martin, G. R. (2012). Extreme binocular vision and a straight bill facilitate tool use in New Caledonian crows. *Nature Communications, 3*, 1–7. http://dx.doi.org/10.1038/ncomms2111

Wood, C. A. (1917). *The fundus oculi of birds, especially as viewed by the ophthalmoscope: A study in comparative anatomy and physiology*. Chicago: The Lakeside Press.

Zentall, T. R., Weaver, J. E., & Clement, T. S. (2004). Pigeons group time intervals according to their relative duration. *Psychonomic Bulletin & Review, 11*, 113–117. http://dx.doi.org/10.3758/BF03206469

Appendix 1

Procedure: instruction

For the humans, the procedure was preceded by the following instruction:

"Hello, thank you very much for your time. I will present you with a set of three objects and you will have only one task: to select one and keep it till the end of the session. We will repeat this twelve times, and every time you are welcome to choose one object and keep it. Once you make your choice, I will present you with the next set. We are recording the session for coding purposes. You remain anonymous and our research team will be the only one to have access to the videos. We are not supposed to communicate during the session, but I will be happy to answer any questions afterwards. Is everything clear?"

Coding: Inter-observer agreement

We aimed to determine the duration of single gazes based on the visual behaviors of the subjects. Thus, we formulated a definition of the single gaze for two independent raters to obey in coding. For each rater, a series of time intervals was generated, and then the inter-observer agreement was determined. Although Cohen's kappa is a common index in such cases, it could not have been applied here. We needed a tool that would estimate the accuracy of the overlap between two interval patterns, generated by the raters for the same recording. This is why we used time-unit kappa, which seemed to serve our purposes best (Bakeman et al., 2009).

It is intuitive that the interval patterns should be quite similar (if the definition is conservative enough), but also the onsets and offsets of the intervals may slightly differ between the raters. Time-unit kappa succeeds in "giving credit for near misses and increasing the magnitude of κ" (1, p. 139), and consequently may provide a more accurate picture of the agreement in our case. Each recording was divided into consecutive one-second intervals, and for each interval a 0-1 response was determined. Occurrence of coding on the rater's part was counted as 1, and its lack was counted as 0. The 0-1 responses for each interval were subsequently assembled into a rater-specific pattern, and finally, an inter-rater kappa coefficient was calculated between these two patterns. There was substantial agreement, with a 0.865 coefficient for the ravens and slightly lower 0.828 for the humans.

All statistical analyses were conducted in *R* (v.3.3.2, the R Foundation for Statistical Computing: http://www.R-project.org). Significance level was set at 0.05.

References

Bakeman, R., Quera, V., & Gnisci, A. (2009). Observer-agreement for timed-event sequential data: A comparison of time-based and event-based algorithms. *Behavioral Research Methods 41,* 137-147. http://dx.doi.org/10.3758/BRM.41.1.137