“Economies of Experience”—Disambiguation of Degraded Stimuli Leads to a Decreased Dispersion of Eye-Movement Patterns

DOI:
10.1111/cogs.12566

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
Cognitive Science

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
“Economies of experience” – disambiguation of degraded stimuli leads to a decreased dispersion of eye-movement patterns

Magdalena Ewa Król¹ and Michał Król²

¹ Faculty of Psychology II, SWPS University of Social Sciences and Humanities in Wrocław, Wrocław, Poland

² Department of Economics, School of Social Sciences, University of Manchester, Manchester, United Kingdom

Corresponding author: Dr. Magdalena Król

Email: mkrol1@swps.edu.pl, Telephone number: 0048 510 763 731

Address: SWPS University of Social Sciences and Humanities, Faculty of Psychology II
Room 58, Ul. Ostrowskiego 30b, 53-238, Wrocław, Poland

This work was supported by the National Science Centre in Poland under Grant 2013/11/D/HS6/04683
ECONOMIES OF EXPERIENCE

Abstract

We demonstrate “economies of experience” in eye-movement patterns, i.e. optimization of eye-movement patterns aimed at more efficient and less costly visual processing, similar to the priming-induced formation of sparser cortical representations or reduced reaction times.

Participants looked at Mooney-type, degraded stimuli that were difficult to recognize without prior experience, but easily recognizable after exposure to their undegraded versions.

As predicted, eye-movement dispersion, velocity and number of fixations decreased with each stimulus presentation. Further analyses showed that this effect was contingent on recognition and the selection of information from the stimulus could be informed by the identity of the presented object.

Finally, our study demonstrates that after exposure to the undegraded version of the stimulus, eye-movement patterns associated with its degraded and undegraded versions become more similar.

This suggests that eye-movement patterns can evolve to facilitate the optimal processing of a given stimulus via experience-driven perceptual learning.

Keywords:
Eye movements, object recognition, top-down, visual attention, perceptual learning, visual processing
1. Introduction

Looking is a skill that, when performed well, allows us to see more of the world - more accurately, quickly and effortlessly. It is a skill of selecting and extracting information from the environment – and the choice of which information to extract and which to ignore is the key to good quality end-product, i.e. what we actually perceive. For example, when looking at ambiguous images that have two possible interpretations (such as the rabbit-duck or the wife-mother in law images), what we perceive depends on the particular features of the image that we attend to (Kawabata & Mori, 1992; Pomplun, Ritter, & Velichkovsky, 1996). As Holm et al. (2008) show, looking at a complex image is akin to detective work, whereby one looks for clues to uncover the image’s identity. Looking in the right places makes it possible to recognize the image more quickly.

Crucially, once the image is recognized, the extensive search for the clues can be terminated. Knowing the areas of the image key to recognition makes it possible to decrease the amount of information extracted from the environment. This saves time and expensive computational resources (Lennie, 2003). All those changes represent “economies of experience”, collectively seen as a knowledge-induced increase in processing efficiency. Economies of experience are also manifested as sparser neural representations (Desimone, 1996; Wiggs & Martin, 1998), or behaviorally as, for example, priming-related increases in accuracy and reduced reaction times (Tulving & Schacter, 1990). In this study, we demonstrate these economies by identifying an experience-related reduction of eye-movement dispersion, i.e. more compact looking patterns resulting from perceptual learning, and other eye-tracking metrics.

At first, eye-tracking would seem suboptimal for the purpose of studying recognition, which generally happens so quickly that eye-movements are often not necessary for it to take place. In humans, eye-movements occur on average every 300 ms (in the range between 50 and 1000 ms) (Henderson & Hollingworth, 1999), while the “gist” of a natural scene can be extracted within a single glance, even if the image is presented for as short as 10 ms, provided that processing of the scene is not disturbed by the subsequent stimulus (Hegdé, 2008). Similarly, in Grill-Spector and Kanwisher (2005) study, all objects presented in the study were identified within 167 ms and most within 50 ms, with basic categorization happening even more quickly. Thorpe, Fize and Marlot (1996) demonstrated that stimulus duration of 20 ms is sufficient for basic identification of the presented object, with 150 ms required for its processing. Finally, DiCarlo and Maunsell (2000) reported that almost 90% of
neural responses in the inferotemporal cortex, responsible for object recognition, were unaffected by the presence of unrestricted eye movements (free viewing), which again suggests that basic object recognition is independent of eye-movements. Nevertheless, longer stimulus presentations that allow for eye-movements naturally lead to more detailed identification of the stimulus (Fei-Fei et al., 2007). Similarly, eye-movements are necessary to identify objects in a scene distributed further than 2-3° visual angle from the center of a fixation (Findlay & Gilchrist, 2003).

Most importantly, investigating eye-movement patterns is an excellent tool to study active aspects of vision, i.e. the strategic selection and extraction of information from the environment based on their importance or task-relevance (Henderson, 2003). Even though bottom-up factors like saliency do influence eye-movement patterns, gaze direction and duration are strongly influenced by top-down factors (Henderson, 2007; Rayner, 2009). In fact, in the process of perceptual learning, eye-movement patterns become increasingly shaped by goals and experience-driven top-down control (Henderson, Weeks, & Hollingworth, 1999). For this reason, eye-tracking is particularly well-suited for the study of top-down optimization of perception.

In order to utilize this potential, we bypassed the limitations related to the application of eye-tracking to recognition by using stimuli similar to Gregory’s (1970) famous Dalmatian picture or Mooney-type stimuli, i.e. highly degraded, black and white images (Mooney, 1956). Those type of stimuli are very difficult or impossible to recognize without prior experience, but once learned, they are easy to interpret and their meaning becomes difficult to ignore. Even single exposure to the disambiguated version of such image is enough to trigger stable, irreversible change in the perception of the degraded stimulus- what Ahissar and Hochstein (2004) call “the Eureka effect”.

Similar images have been used in the study of neural correlates of object recognition (Bertrand & Tallon-Baudry, 2000; Dolan et al., 1997; Hsieh, Vul, & Kanwisher, 2010; Landis, Lehmann, Mita, & Skrandies, 1984; Pegna, Khateb, Michel, & Landis, 2004). The advantage of these type of stimuli is that they allow us to compare recognition to the absence of recognition, without necessitating any changes in the physical properties or appearance of the stimulus. Additionally, that means that any changes in the perception of the stimulus must be caused purely by top-down factors, such as perceptual learning.
Another advantage of these type of stimuli is that they avoid the aforementioned problem of high-speed recognition of natural images, relative to the intrinsic limits to the eye movements’ velocity. They are sufficiently difficult to prevent quick recognition, and very often, without prior experience with the disambiguated versions - any recognition at all. Each stimulus is a perceptual puzzle that has to be approached in the right way to reveal its meaning- and we hypothesize that the right approach is linked to the eye movement patterns. The purpose of the process of moving and fixating the eyes on selected areas of the visual field is to choose the right type of information from the environment to perform the task at hand (Buswell, 1935; Yarbus, 1967). Recent research has shown that cognitively complex tasks (such as estimating the material circumstance of people depicted in the picture) are related to distinct patterns of eye movements, optimized to the given task (Borji & Itti, 2014). By using highly degraded stimuli that delay and obstruct recognition, the process of extracting information from the environment to solve the “perceptual puzzle” can be observed in “slow motion”. As Noton and Stark (1971) argued, presenting stimuli in difficult viewing conditions forces the observer to process the image feature by feature, thereby revealing the strategy behind its processing.

We hypothesized that recognition, compared to its absence, will be accompanied by distinct patterns of eye-movements that optimally sample the environment for information necessary for visual processing. It might appear at first that the types of stimuli we use are artificial and therefore eye-movement correlates of their recognition are not applicable to the mechanism of recognition in natural conditions. However, as Hegdé and Kersten (2010) argue, ambiguity is inherent in natural images and experience-based disambiguation is a basic function of perception, even if it usually seems to happen effortlessly and rapidly. This ambiguity often stems from the presence of noise in the visual input, defined here broadly as all information irrelevant to the task. Thus, one of the purposes of recognition is the detection of signal information amid noise (Ahissar & Hochstein, 2004). Studying eye movements related to that process can therefore show how the visual system selects information that enables the proverbial separation of the wheat from the chaff. We hypothesize that this process will be optimized by perceptual learning resulting from exposure to a disambiguated version of the stimulus. Furthermore, prior experience with the disambiguated stimuli should bring about a change in prospective eye-movement patterns, optimizing them to perceive that particular stimulus. Specifically, we hypothesize that this change will be accompanied by a decrease in the amount of information extracted from the environment, measured by the decreased
dispersion of eye movements. In other words, recognition, compared to the absence of recognition, should be related to a more compact looking pattern.

The idea that sparsity follows experience is epitomized in repetition suppression, that is the reduction of neuronal activity that follows repetition of the stimulus. According to Desimone (1996), it is caused by a “sharpening” of neuronal representation, that is a decrease in the number of neurons participating in the neural response to the stimulus. Neurons that are not necessary for the representation of the stimulus cease to respond with repeated exposure. Wiggs and Martin (1998) went even further, by hypothesizing that repetition of the stimulus leads to a neural representation that does not reflect the complete stimulus, but only those features necessary for its identification. This not only allows for an even sparser, more economical representation, but may also facilitate future recognition by increasing the tolerance to inevitable changes to the object appearance (Kourtzi, Dicarlo, Glimcher, & Kanwisher, 2006). The increase in repetition-induced efficiency of neural response is also reflected in improved behavioral markers of performance, such as higher accuracy or decreased response time, usually referred to as repetition priming (Tulving & Schacter, 1990). However, this process is unlikely to take place (or will occur to a limited extent) if recognition does not occur (Henson, Shallice, & Dolan, 2000). Both repetition priming and suppression are related to experience-induced optimization of neural representation – if there is no pre-existing neural representation of the presented object, neither recognition nor repetition priming or suppression will take place. In such case, stimulus repetition will lead to the formation of a new neural assembly, and not its optimization (Gruber & Müller, 2005). Thus, it would seem that recognition (that implies prior existence of presented object’s neural representation) is a prerequisite for the economies of experience - i.e. for sparser neural representations, shorter processing times, higher accuracy and – as we hypothesize – more compact eye-movement patterns.

Existing eye-tracking studies demonstrate the general idea of experience-related economies reflected in eye-tracking measures. According to the information reduction hypothesis (Haider & Frensch, 1999a, 1999b), experience leads to reduced allocation of attention to irrelevant information. Similar results were obtained from many studies comparing eye-movement patterns of experts and novices. The common finding is that experts have more fixations on task-relevant areas than novices (for a meta-analysis, see: Gegenfurtner, Lehtinen, & Säljö, 2011). Similarly, according to the adaptive scanning hypothesis (Myers & Gray, 2010), the number of fixations per stimulus decreases with stimulus repetition. For
example, Hidalgo-Sotelo, Oliva and Torralba (2005) found that presence of a strong prior led to a decrease in gaze duration on the target. In their classic study, Noton and Stark (1971) demonstrated that the number and the average duration of fixations to a poorly-lit stimulus decreased in the recognition phase compared to the earlier learning phase. Hannula et al. report a decrease in the number of fixation for old compared to novel stimuli (2010; 2012). Holm et al. (2008) show a gradual “homing in” of fixations to the target-related regions of perceptually-demanding stimuli just before recognition is announced by the participant. Finally, Over et al. (2007) found a gradual decrease in saccade amplitude coupled with an increase in fixation duration in a visual search task, as the participant was “homing in” on the target.

We therefore hypothesize that degraded stimulus repetition after disambiguation will lead to a decrease in eye-movement dispersion, but only for those stimuli that were successfully recognized. We used eye-movement dispersion as an estimate of the amount of information extracted from the stimulus. We hypothesized that perceptual learning would lead to a decrease in the number of extracted information as a result of allocating attention to those areas of the stimulus that are necessary for recognition.

We also hypothesize that the ability to look “the right way” at the degraded stimuli will correlate with recognition. Noton and Stark (1971) found idiosyncratic scanpaths that were stable across repetitions of the same stimulus presented to the same observer but changed with the change of stimulus or observer. According to Myers and Gray (2010), such stable stimulus-related scanpaths can be seen as perceptual skills, i.e. acquired units of saccade sequences applied to optimally process given stimulus and matched to its internal representation (Noton & Stark, 1971). According to the adaptive scanning hypothesis (Myers & Gray, 2010), apart from the decrease in the number of fixations per stimulus, we should observe the scanpaths of the same repeated stimulus becoming more similar with experience, reflecting the adoption of the scanpath to the optimal way of processing that particular image. Thus, we hypothesize that the degree of similarity between eye-movement patterns accompanying the disambiguated vs. degraded stimuli should be correlated with higher rates of recognition of the latter. That is, observers who look at the degraded stimuli similarly to how they look at its disambiguated version would have a higher chance of recognizing the degraded stimulus.

To summarize, we hypothesized that 1) perceptual learning resulting from exposure to an undegraded stimulus will lead to a decreased eye-movement dispersion, i.e. more compact
looking patterns. The reason for this reduction is the knowledge-driven selection of those features of the image that would allow for a successful recognition of the object. Because of that, we further hypothesized that 2) the reduction in eye-movement dispersion will only occur in the case of recognized stimuli. We also hypothesized that 3) presentations of the same stimulus eliciting the same behavioral response (either recognized or unrecognized) will share more similar eye-movement patterns compared to presentations followed by mismatched responses. In particular, recognition of degraded stimuli will correlate with the similarity of the associated eye-movement patterns to the “optimal” patterns accompanying the disambiguated stimuli.

To test these hypotheses, we selected Mooney-type, degraded stimuli that were difficult to recognize without prior experience, but after disambiguation should be readily identified. Stimuli were obtained by transforming photographs representing natural scenes. Each stimulus was displayed three times in direct succession: firstly in the degraded version (initial presentation), then in the original, undegraded version (undegraded presentation), then again in the degraded version (immediate-after-undegraded presentation). The fourth and final presentation (delayed presentation) of the stimulus (in the degraded version) was delayed and interspersed by different stimuli. The purpose of the second (undegraded) and third (degraded) presentations was to enable perceptual learning. The initial presentation of the degraded stimulus was our measure of perception without prior experience of the stimulus. The last presentation of the degraded stimulus (delayed presentation) was our measure of perception with prior experience. This presentation was delayed and interspersed with the presentation of other stimuli in order to make sure that recognition was a result of re-activation of the relevant representation, and not resulting from holding an active representation of the previous, undegraded stimulus.

2. Method
2.1 Pilot study
We wanted to find stimuli that would be very difficult to recognize without prior experience, but once recognized, would be easily interpretable in the future (we will call this the Mooney property). For that purpose, we conducted a pilot study with the aim of obtaining ratings of how difficult it was to recognize degraded images before and after seeing their undegraded versions.
ECONOMIES OF EXPERIENCE

The stimuli were 440 high-quality photographs of natural scenes containing objects or people, purchased from Dreamstime (http://www.dreamstime.com/). Half of the photographs contained at least one person (social stimuli) and in the remaining 220 photographs there were no people, but instead, they presented animals, plants, food or man-made objects. All photographs were in landscape-orientation. Photographs were processed using Adobe Photoshop CS2 to remove color information (grayscale mode) and to adjust their size to fit the screen of 1280x720, subtending 15.9° x 27.7° of visual angle. Photographs prepared in this way were used as undegraded stimuli. Next, each photograph was additionally processed with a “stamp” filter, which allows obtaining Mooney-type images. A similar effect can be obtained by removing high-frequency spatial information from the image, followed by an application of a very high contrast, effectively converting grays into black and white.

The pilot study was designed and performed in a very similar manner to the main study. In the pilot study, there were 11 sessions with 10 images each. Participants were shown each image four times. The first three presentations were consecutive and images were presented in the following order: degraded image (initial presentation), undegraded image, degraded image (immediate-after-undegraded presentation). Finally, the degraded image was presented one more time, but after all other stimuli in the session were presented three times (delayed presentation). After each presentation participants were requested to indicate whether they recognized the image, by pressing a key.

The pilot study was performed on 63 (25 male) participants, mean age= 25.3 (SD=3.92). Data obtained in the pilot were used in the stimuli selection procedure described in the next section and in stimuli selection for other studies.

2.2 Main study

2.2.1 Stimuli

For each image and for each of its four presentations in the pilot study, we calculated the percentage of participants who indicated that they recognized the image in that particular presentation. We excluded all pictures that were recognized by more than 75% of participants in the initial presentation (too easy-criterion 1) and all pictures that were recognized by less than 75% of participants in presentation 4 (too difficult-criterion 2). Next, we subtracted the percentage of participants who recognized a particular image in the initial presentation from the percentage of participants who recognized it in the final, delayed presentation. Then we excluded all photographs where that difference was smaller than 50% (does not possess the Mooney property-criterion 3). Finally, from the remaining photographs, we chose two sets
ECONOMIES OF EXPERIENCE

of 40 social and 40 non-social images that did not differ significantly in terms of either of the three criteria. For examples of stimuli used in the study, see Fig. 1.

Fig. 1 about here

2.2.2 Participants
Participants were 148 (105 females) volunteers, aged between 18-46 ($M = 23.7; SD = 6$). Participants took part in the study in exchange for $7 per hour. All participants had normal or corrected to normal eyesight. The study procedure was approved by the SWPS University of Social Sciences and Humanities, Faculty of Psychology II in Wroclaw Research Ethics Committee, in accordance with the Helsinki Declaration.

2.2.3 Procedure
Participants’ eye movements were recorded using a remote eye-tracking device SMI RED250Mobile, with a sampling rate of 60 Hz and gaze position accuracy of 0.4°. Participants were seated 70 cm from the computer screen. The experiment was programmed in C#. Participants completed a 5-point calibration and 4-point validation in-house procedure. Additionally, the in-house software periodically checked the quality of the acquired data and if it fell below a predefined criterion, participant repositioning was prompted after the end of the session.

The experiment consisted of 8 sessions. Each session consisted of 10 stimuli, displayed four times each in a specific order. Each stimulus appeared only four times in the whole study and always within the same session. Half of the stimuli within each session were social and half were non-social. The assignment of stimuli to the session and their order within each session were randomized for each participant.

Within every session, each of the 10 stimuli was displayed three times consecutively – first in the degraded version (initial presentation), then in the undegraded version and, immediately after that, again in the degraded version (immediate-after-undegraded presentation). Once each of the 10 stimuli was displayed thrice in this way, all 10 were successively displayed one last time in the degraded version in a random order (delayed presentation). Therefore, each session consisted of 40 trials.

For each of the four presentations of an image, the trial began with a fixation cross on white background displayed for 300ms, followed by the stimulus presented for 1000 ms. Next, participants were requested to respond to a question “Have you recognized what this image
represents?” by pressing YES or NO key (labeled on the keyboard). The trial was concluded with a blank screen displayed for 500 ms. For a schematic diagram of a session see Fig. 2.

Fig. 2 about here

We used relatively short presentation times (1000 ms) to capture only early recognition - we wanted to test a large set of photographs to be able to generalize the findings and, given four presentations of each stimulus, the task was quite tiring. Because of their intrinsic difficulty, some of the degraded stimuli that were not immediately recognized could sometimes be recognized after prolonged study. However, in those cases, early eye-movements would be similar to the eye-movements typical for unrecognized images, while behaviorally recognition would be indicated, potentially introducing noise to the results. For this reason, we limited presentation time to allow for a clear distinction between stimuli recognized within 1000 ms, and those that were not recognized within that time-frame.

3. Results

3.1 Data analysis

Reaction times were measured from the onset of the stimulus to the key press. The proportion of recognized stimuli was measured as the proportion of ‘Yes’ responses to the question ‘Have you recognized what this image represents?’ Fixations were identified using the SMI Event Detector software, with minimum fixation length of 80 ms.

Moreover, the purpose of our analysis was to obtain an approximation of the size of the area scanned by eye-movements, by estimating whether eye-movements were ‘dispersed’ or concentrated in a small number of clusters. Because subjects had to process a relatively large amount of visual information during the short duration of the trial, the number of identified fixations per trial was small (M=1.72, SD=0.43).

Given such a low number of fixations, it is impossible to reliably compare experimental conditions in terms of dispersion measures that use fixation data as input (e.g. 'dispersion' calculated for a set of two fixation points is not informative). For this reason, in the case of measuring dispersion, we analyzed the raw data samples directly. According to Holmqvist (2011, p. 357), “These x-y positions are the most reliable and detailed position data, [... whereas] fixation positions are additionally subject to the peculiarities of filters, the selected fixation detection algorithm and its settings.” In particular, analyzing the eye positions stored in raw data samples, often called 'points-of-regard' (PORs), allows for a much greater
temporal resolution compared with fixation analysis, which is essential given the short
duration of our trials.

Accordingly, we use eye metrics that are typically employed to process raw sample data prior
to identifying fixations. In particular, the mean distance between two successive PORs is
listed by Holmqvist (2011, p. 155) as 'one of the varieties of dispersion measures', and closely
related to the 'root mean square distance' (again, between two successive PORs) which is
commonly used as a dispersion measure. Nevertheless, when expressed in degrees of visual
angle per second, the mean distance between successive PORs is more often referred to as
(average) 'eye-velocity'. For instance, 'velocity-based' event detection algorithms split PORs
into saccades or fixations depending on whether the velocity is above or below a specified
threshold (Salvucci & Goldberg, 2000). Although the outcome of this process is sensitive to
how the threshold is set (Shic, Scassellati, & Chawarska, 2008) by using raw samples we can
directly report velocity in any part of a trial without the need to specify the threshold and then
working with the detected event data.

Velocity was thus calculated as the average Euclidean distance between two subsequent
PORs. We excluded those samples that were flagged as missing data by the eye-tracker and
eliminated all trials with more than 50% of such samples.

Another measure we used, in this case associated exclusively with dispersion, is 'spatial
coverage', defined as the fraction of cells of a 16:9 grid (separating the screen into 80x80
pixel tiles) that contain at least one POR (Holmqvist et al., 2011, p. 367). To avoid confusion,
we will henceforth refer to the mean distance between successive PORs as 'velocity', and to
the fraction of cells covered by the PORs as 'dispersion'. We analyzed the data using several
other dispersion measures and the results were very similar irrespective of the measure,
which means that the patterns in the data described here are robust to changing the calculated
eye-metrics.

Two participants were excluded based on extremely high velocity values (more than 10 times
the average velocity), suggesting a technical malfunction.

There were no significant differences between social and nonsocial stimuli for any of the
above measures, and for this reason, we collapsed the data across this variable.
To mitigate the effect of outliers, all values higher or lower than mean and two standard deviations were replaced with the mean +/- two standard deviations. The percentage of cases replaced in this way varied between 3.7-5% depending on the measure.

Additionally, to normalize the data and allow parametric statistical analyses, we used a cube root transformation on reaction times and velocity measure (all figures represent back-transformed data). We used nonparametric tests for recognition rates, due to near-ceiling effect after the first presentation, which rendered the data essentially non-parametric, as well as for the similarity measures, where the deviation from normality was too large to be corrected. All other measures were normally-distributed.

### 3.2 The effect of presentation

#### 3.2.1 Proportion of recognized images

We performed a Friedman test to analyze the proportion of recognized stimuli to detect differences between the four presentations of the stimulus (initial, undegraded, immediate-after-undegraded, delayed). The test was followed by a post hoc analysis using Bonferroni-corrected Wilcoxon signed-rank tests. Post-hocs were performed in the fashion of repeated contrasts, resulting in three tests per analysis and a corrected alpha level of 0.017.

The main effect of presentation was significant, $\chi^2(3)=295.74$, $p<.001$. Post-hocs revealed that the initial presentation had lower recognition rates than the undegraded presentation, $Z=-10.43$, $p<.001$, $r=-.86$. The undegraded presentation, in turn, had significantly higher recognition rates than the immediate-after-undegraded presentation, $Z=-3.99$, $p<.001$, $r=-.33$. Finally, recognition rates were significantly lower for the delayed presentation compared to the immediate-after-undegraded presentation, $Z=-4.20$, $p<.001$, $r=.35$ (Fig. 3a).

#### 3.2.2 Reaction times

We performed a one-way repeated-measures ANOVA with presentation (four levels: initial, undegraded, immediate-after-undegraded and delayed) as the factor. There was a significant main effect of presentation, $F(3,435)=181.03$, $p<.001$, $\eta_p^2=.56$.

Contrasts revealed that reaction times were not significantly different in the undegraded presentation, compared to the initial presentation, $F(1,145)=3.49$, $p=.06$, $\eta_p^2=.02$. However, reaction times were significantly shorter in the immediately-after-undegraded presentation, compared to the undegraded presentation, $F(1,145)=56.85$, $p<.001$, $\eta_p^2=.28$. Finally, reaction times were significantly shorter in the delayed presentation, compared to the immediately-after-undegraded presentation, $F(1,145)=151.73$, $p<.001$, $\eta_p^2=.51$ (Fig. 3b).


3.2.3 Average number of fixations
We performed a one-way repeated measures ANOVA with presentation (four levels: initial, undegraded, immediate-after-undegraded and delayed) as the factor. There was a significant main effect of presentation, $F(3,435)=119.87$, $p<.001$, $\eta^2_p=.45$.

Contrasts revealed that the average number of fixations was significantly lower in the undegraded presentation, compared to the initial presentation, $F(1,145)=43.67$, $p<.001$, $\eta^2_p=.23$. In turn, there were on average fewer fixations in the immediately-after-undegraded presentation compared to the undegraded presentation, $F(1,145)=114.56$, $p<.001$, $\eta^2_p=.44$. However, the number of fixations was significantly higher in the delayed presentation, compared to the immediately-after-undegraded presentation, $F(1,145)=17.03$, $p<.001$, $\eta^2_p=.11$ (Fig. 3c).

3.2.4 Average duration of fixations
We performed a one-way repeated measures ANOVA with presentation (four levels: initial, undegraded, immediate-after-undegraded and delayed) as the factor. There was a significant main effect of presentation, $F(3,435)=26.42$, $p<.001$, $\eta^2_p=.15$.

Contrasts revealed that the average duration of fixations was significantly longer in the undegraded presentation, compared to the initial presentation, $F(1,145)=41.73$, $p<.001$, $\eta^2_p=.22$. There was no significant difference between the immediately-after-undegraded presentation and the undegraded presentation, $F(1,145)=3.44$, $p=.07$, $\eta^2_p=.02$. However, the duration of fixations was significantly lower in the delayed presentation, compared to the immediately-after-undegraded presentation, $F(1,145)=15.86$, $p<.001$, $\eta^2_p=.10$ (Fig. 3c).

3.2.5 Eye movement dispersion
We performed a one-way repeated measures ANOVA with presentation (four levels: initial, undegraded, immediate-after-undegraded and delayed) as the factor. There was a significant main effect of presentation, $F(3,435)=113.07$, $p<.001$, $\eta^2_p=.44$.

Contrasts revealed that dispersion was significantly lower in the undegraded presentation, compared to the initial presentation, $F(1,145)=76.70$, $p<.001$, $\eta^2_p=.35$. In turn, dispersion was significantly lower in the immediately-after-undegraded presentation, compared to the undegraded presentation, $F(1,145)=55.47$, $p<.001$, $\eta^2_p=.28$. However, there was no significant difference in dispersion between the delayed presentation and the immediately-after-undegraded presentation, $F(1,145)=2.65$, $p=.11$, $\eta^2_p<.01$ (Fig. 3c, see also Fig. 4).
3.2.6 Eye movement velocity
We performed a one-way repeated-measures ANOVA with presentation (four levels: initial, undegraded, immediate-after-undegraded and delayed) as the factor. There was a significant main effect of presentation, $F(3,435)=93.39$, $p<.001$, $\eta_p^2=.39$.

Contrasts revealed that velocity was significantly lower in the undegraded presentation, compared to the initial presentation, $F(1,145)=45.39$, $p<.001$, $\eta_p^2=.24$. In turn, velocity was significantly lower in the immediately-after-undegraded presentation, compared to the undegraded presentation, $F(1,145)=51.44$, $p<.001$, $\eta_p^2=.26$. However, there was no significant difference in velocity between the delayed presentation and the immediately-after-undegraded presentation, $F(1,145)=0.21$, $p=.65$, $\eta_p^2<.01$ (Fig. 3c).

3.3 The effect of recognition
In order to analyze the effect of recognition, we broke down dispersion and reaction times values into two groups: corresponding to “Yes” and “No” responses. In undegraded and immediate-after-undegraded presentations the number of “Yes” responses was very high (95% and 93%). For this reason, this analysis was only performed on the initial (30% of recognized images) and delayed (91% of recognized images) presentations. It is also important to point out that 12 participants recognized all images in the delayed presentation which resulted in missing data points in this analysis. For all of the analyses below, we performed a 2 (presentation: initial vs delayed) x 2 (recognition: not recognized vs recognized) repeated-measures ANOVA. Significant interactions were followed by Bonferroni-corrected post-hoc t-tests, with alpha=.0125, given four comparisons we performed for each measure.

3.3.1 Reaction Times
The initial presentation was related to significantly longer reaction times than delayed presentation, $F(1,133)=148.81$, $p<.001$, $\eta_p^2=.53$. Recognized images were related to significantly shorter reaction times, $F(1,133)=91.61$, $p<.001$, $\eta_p^2=.43$. The interaction between presentation and recognition was also significant, $F(1,133)=82.81$, $p<.001$, $\eta_p^2=.38$, which means that the difference in reaction times between recognized and not recognized images was significantly larger for the delayed presentation, compared to the initial presentation. Post-hocs revealed that the difference in reaction times between recognized and
not recognized images was not significant for the initial presentation, $t(133)=0.35$, $p=.73$, $d=0.03$, but it was significant for the delayed presentation, $t(133)=12.54$, $p<.001$, $d=1.08$. Additionally, the difference in reaction times between the initial and delayed presentation was not significant for images that were not recognized ($t(133)=0.18$, $p=.86$, $d=0.02$, but it was significant for recognized images ($t(133)=14.58$, $p<.001$, $d=1.26$ (Fig. 5a).

### 3.3.2 Dispersion

There was a significant effect of presentation ($F(1,133)=60.15$, $p<.001$, $\eta^2_p=.31$)- the delayed presentation was related to smaller dispersion, compared to the initial presentation. The main effect of recognition was also significant, ($F(1,133)=29.61$, $p<.001$, $\eta^2_p=.18$)- recognized images were related to smaller dispersion than images that were not recognized.

The interaction between presentation and recognition was also significant, $F(1,133)=34.06$, $p<.001$, $\eta^2_p=.20$. Post-hocs revealed that the difference between recognized and not recognized images was not significant for the initial presentation, $t(133)=0.28$, $p=.78$, $d=0.07$. However, in the delayed presentation, recognized images, compared to unrecognized images, were related to smaller dispersion, $t(133)=7.02$, $p<.001$, $d=0.61$. The difference in dispersion between the initial and delayed presentations was significant for the recognized images ($t(133)=10.68$, $p<.001$, $d=0.92$), but not for the images that were not recognized ($t(133)=1.71$, $p=.09$, $d=0.15$ (Fig. 5b).

### 3.3.3 The average number of fixations

Delayed presentation, compared to the initial presentation, was related to a lower average number of fixations ($F(1,133)=13.46$, $p<.001$, $\eta^2_p=.09$). Recognized images, compared to unrecognized images, were related to a lower average number of fixations ($F(1,133)=27.97$, $p<.001$, $\eta^2_p=.17$). The interaction between recognition and presentation was also significant ($F(1,133)=19.59$, $p<.001$, $\eta^2_p=.13$).

Post-hocs revealed that the difference between recognized and not recognized images was not significant for the initial presentation, $t(133)=0.91$, $p=.36$, $d=0.08$. However, in the delayed presentation, recognized images, compared to unrecognized images, were related to fewer fixations ($t(133)=6.25$, $p<.001$, $d=0.71$). The difference in the number of fixations between the initial and delayed presentations was significant for the recognized images ($t(133)=8.16$, $p<.001$, $d=0.71$), but not for the images that were not recognized ($t(133)=2.09$, $p=.04$, $d=0.18$ (Fig. 5c).
ECONOMIES OF EXPERIENCE

3.3.4 The average duration of fixations
Delayed presentation, compared to the initial presentation, was related to a shorter duration of fixations (F(1,133)=45.28, p<.001, $\eta^2_p=.25$). There was no effect of recognition, F(1,133)=0.94, p=.34, $\eta^2_p=.01$, and no significant interaction between recognition and presentation, F(1,133)=0.002, p=.96, $\eta^2_p<.01$ (Fig. 5d).

Fig. 5 about here

3.3.5 Eye-movement velocity in the early and late period of stimulus presentation
To investigate the effect of time on dispersion, we calculated eye movement velocity for each half of the image presentation time. Next, we performed a 2 (presentation: initial vs delayed) x 2 (recognition: not recognized vs recognized) x 2 (time: early vs late) repeated-measures ANOVA.

Velocity was significantly higher for the initial presentation, compared to the delayed presentation, F(1,133)=61.58, p<.001, $\eta^2_p=.32$. Recognized images were related to significantly lower velocity than unrecognized images, F(1,133)=15.14, p<.001, $\eta^2_p=.10$. Velocity was significantly lower in the first 500 ms, compared to the last 500 ms of image presentation, F(1,133)=568.25, p<.001, $\eta^2_p=.81$. The interaction between time and presentation was not significant, F(1,133)=2.62, p=.11, $\eta^2_p=.02$. However, the interaction between presentation and recognition was significant, F(1,133)=7.01, p=.01, $\eta^2_p=.05$, which means that the difference between velocity for recognized and not recognized images was larger in the delayed presentation, compared to the initial presentation. Moreover, the interaction between recognition and time was also significant, F(1,133)=62.54, p<.001, $\eta^2_p=.32$, which means that the difference between recognized and not recognized images was larger in the last 500 ms of image presentation than in the first 500 ms. Finally, there was a significant three-way interaction between recognition, time and presentation, F(1,133)=32.13, p<.001, $\eta^2_p=.20$.

To understand the three-way interaction, the main analysis was followed by four Bonferroni-corrected post hoc t-tests, which revealed that for the initial presentation, the difference in velocity between recognized and not recognized images was not statistically significant, both in the early period of image presentation, t(133)=.62, p=.54, d=.05, and in the late period: t(133)=2.36, p=.02. d=.20. However, in the delayed presentation, velocity was significantly higher for recognized images, compared to unrecognized images in the early period,
t(133)=3.49, p=.001, d=0.30, while in the late period, velocity was significantly lower for recognized images, compared to unrecognized images, t(133)=7.76, p<.001, d=0.67 (Fig. 6).

3.3.6 Deceleration onset
For each trial, the eye deceleration onset is calculated as the smallest trial duration percentage $p$, such that the maximum Euclidean distance between any two subsequent eye positions recorded during the first $p\%$ of the trial is higher than the analogous maximum distance value in the remaining part of the trial. In other words, if the eye ceases moving rapidly and settles on a given point early on in the trial, the eye deceleration onset will be close to zero. In contrast, if eye movements continue to accelerate throughout the trial, it will be close to 100%.

The main effect of presentation was insignificant, $F(1,133)=3.18$, $p=.08$, $\eta^2_p=.02$. However, the main effect of recognition was significant, $F(1,133)=97.45$, $\eta^2_p=.42$. In general, the eye decelerated earlier for recognized images (at $M=59\%$ of image presentation, 95% CI [58%, 60%]) than for unrecognized images (M=65%, 95% CI=[63%, 66%]). The interaction between presentation and recognition was also significant, $F(1,133)=40.11$, $p<.001$, $\eta^2_p=.23$, which means that the difference between recognized and not recognized images was larger in the delayed presentation, compared to the initial presentation. The onset of deceleration can be observed in Fig. 7, which represents the dynamics of eye-movement deceleration over all samples within the trial.

Fig. 7 about here

3.4 Similarity between eye-movement patterns for degraded and undegraded images
We conducted pairwise comparisons between eye-movement patterns associated with different presentations of the same image, by calculating the time warping distance (Berndt & Clifford, 1994) between the sequences of eye-positions corresponding to the compared expositions, as a measure of the eye-movement pattern dissimilarity between the two images. Recall that each sequence contained 60 pairs of X-Y coordinates, one for each eye (the stimulus presentation time was 1s. and the sampling rate 60Hz).

We compared the levels of (dis)similarity between 1) initial and undegraded presentations, 2) initial and delayed presentations and 3) undegraded and delayed presentations (Fig. 8). There was a significant difference between the three comparisons of dissimilarity, $\chi^2(2)=173.1$, $p<.001$. Wilcoxon signed-rank tests (with a Bonferroni-corrected alpha level equal to 0.025) showed that eye-movement patterns’ similarity between the delayed and the undegraded
Presentations was higher than between the initial and undegraded presentation, Z=-9.43, p<.001, r=-.55. Additionally, the eye-movement patterns’ similarity between the initial and delayed presentations was higher than between delayed and undegraded presentations, Z=-5.88, p<.001, r=-.34.

Next, we repeated the analysis separately for “same responses” cases, where both of the compared presentations were followed by an identical participant response (i.e. both presentations were recognized, or both unrecognized) and for “different responses” cases, where exactly one presentation was recognized. It is important to note that most cases of “same responses were in fact cases of double recognitions because recognition rates were equal to 95% in case of undegraded and 91% in case of delayed presentation.

Wilcoxon signed-rank tests (with Bonferroni-corrected alpha level equal to 0.017) showed that “same responses” were related to more similar eye-movement patterns than “different responses” for all three comparisons: between initial and undegraded presentations, Z=-3.6, p<.001, r=-.21, between delayed and undegraded presentations, Z=-5.01, p<.001, r=-.34, and between initial and delayed presentations, Z=-2.8, p<.001, r=-.29 (Fig. 8).

4. Discussion

The purpose of the study was to test three hypotheses. Firstly, we proposed that exposure to the undegraded version of a Mooney-type image will result in decreased eye-movement dispersion. This would result from an experience-driven reduction in the amount of information extracted from the environment, thereby being a signature of optimization of the processing of the stimulus. Secondly, we hypothesized that this reduction of the eye-movement dispersion will occur only when the stimulus was recognized. While the first hypothesis was confirmed and each subsequent presentation of the stimulus was associated with a decrease in the eye-movement dispersion, this reduction applied to a smaller extent also to cases where recognition of the stimulus was not achieved.

We performed some additional exploratory analyses of eye-movement patterns that lent additional support to our hypotheses. For all measures we used, there was a clear effect of experience. For the majority of measures (with the exception of average duration of fixations), there was a significant effect of recognition.
Finally, we confirmed our third hypothesis, that a higher degree of similarity of eye-movements between degraded and undegraded stimuli would be related to higher recognition rates, reflecting more efficient looking patterns.

4.1 Methodological considerations

The first observation we would like to make is that our stimuli did possess the Mooney characteristics. On average, participants reported recognizing 30% images in the initial presentation. For the undegraded presentation and the presentations that follow it, the percentage rises to over 90%. That means that most stimuli were very difficult to interpret without prior experience, but once disambiguated, they became easily interpretable.

In our study, we presented each stimulus four times, where the second presentation revealed the original, undegraded image, while in the remaining three presentations the degraded version was displayed. The first presentation (initial) was our measure of perception pre-learning, while the last (delayed) presentation was our measure of post-learning perception. The purpose of both the second (undegraded) and third (immediate-after-undegraded) presentations were to induce learning. In some other studies using similar paradigms (Dolan et al., 1997; Loth, Gómez, & Happé, 2010) perceptual learning was limited to the presentation of the disambiguated version of the stimulus. In Pegna et al. (2004) study, participants were additionally given paper printouts of the undegraded stimuli that they were able to study. In our experiment, the learning phase consisted of the presentation of both undegraded and degraded stimuli in succession, to make sure that participants learn to see the similarity between the two.

We observed a small but significant decrease in recognition rates from the undegraded (95% recognized) to immediate-after-undegraded presentation (93% recognized), and again, a loss of recognition from 93% to 91% between immediate-after-undegraded and delayed presentations. The loss of recognition between undegraded and immediate-after-undegraded presentations suggests difficulty in applying the image interpretation imposed by the undegraded stimulus onto the degraded image. The loss of recognition between the immediate-after-undegraded and delayed presentation probably resulted from the imperfect re-activation of the learned percept, i.e. from imperfect perceptual learning. Without the immediate-after-undegraded presentation, the loss of recognition between the undegraded and delayed presentations might have been significantly larger. Overall, this demonstrates that
immediate-after-delayed presentation was helpful in inducing learning of the degraded stimulus interpretation.

4.2 Recognition versus optimization of recognition

Recognition rates were the only measure that deteriorated slightly after the undegraded presentations. Moreover, recognition rates improved rapidly between the initial and undegraded presentation, and then seemed to reach the ceiling by remaining at a similar, but slightly lower level throughout the rest of the presentations.

In contrast, reaction times significantly improved with each presentation. The only exception was no significant improvement in reaction times between the initial and undegraded presentations. This lack of improvement is probably caused by the main burden of recognition falling on the undegraded presentation, and as a result, the demand for processing time was similar to the initial presentation. But apart from this exception, reaction times continued to improve with each subsequent presentation, even as recognition was declining.

All eye-tracking measures displayed a similar pattern- an improvement after each subsequent presentation, up until the delayed presentation, where either lack of further improvement (eye-movement velocity and dispersion) or deterioration (number and duration of fixations) were observed. This stall of progress represents perceptual memory loss caused by the delay between the presentations. These results are consistent with the previous findings regarding the interpretation of these eye-tracking measures. For example, a higher number of fixations is related to difficulty in interpreting the visual stimulus (Ehmke & Wilson, 2007; Rötting, 2001). In contrast, fewer fixations are a sign of expertise (for example Reingold, Charness, Pomplun, & Stampe, 2001; Schoonahd, Gould, & Miller, 1973). Moreover, longer duration of fixations signifies both deeper processing (Henderson et al., 1999; Mackworth & Bruner, 1970) and expertise (Reingold et al., 2001; Savelsbergh, Williams, Kamp, & Ward, 2002).

However, what is important here is that both reaction times and all eye-movement measures improved between undistorted presentation (where usually recognition occurred) and the subsequent distorted presentation (immediate-after-undistorted presentation). This means that the improvement could not be related to recognition, as recognition between these two presentations declined.

While the most important goal of visual processing is naturally to achieve recognition, the secondary goal is to achieve it within the shortest time and at minimal cost and effort. Our
ECONOMIES OF EXPERIENCE

results suggest that eye-movement dispersion is more strongly related to the process of optimization of information processing than to recognition per se.

Quick recognition allows for quicker action, while economic computation makes it possible to save computational resources (Lennie, 2003) and bypass the inherent limits of the nervous system (Barlow, 1981). There is, of course, a degree of unavoidable trade-off between accuracy of recognition and cost/time minimization – and on many occasions, accuracy may be sacrificed to a certain extent with the goal of decreasing speed or costs related to processing. An example of this is “change blindness” - representation of a scene is updated only when attention is focused on the area of change, i.e. only if it is deemed important by the attentional system (Rensink, O’Regan, & Clark, 1997).

These two goals - recognition accuracy and its cost/time optimization, though naturally related, are separate, and the optimization process does not end with the achievement of recognition. With each successive presentation, even though recognition did not improve after reaching a close-to-maximum level, both reaction times and eye-movement dispersion continued to decrease, reflecting leaner processing and more focused information search. The same process applies to the “sharpening” of neural representations (Desimone, 1996; Gruber & Müller, 2005; Wiggs & Martin, 1998), where the formation of neural representation continues after recognition, with the purpose of economizing the representation (Kourtzi et al., 2006).

This shows that recognition is not the end of the road in visual processing, and once it is achieved there is still work to do on the optimization of both stimulus related processing and its neural representation.

4.3 Recognition drives optimization of visual processing

We compared reaction times and eye-movement patterns in the first (initial) and the last (delayed) presentations, separately for recognized and not recognized images, to tease out the effect of recognition. We found that improvements in reaction times and eye-movements were driven solely by recognition. That is, with the exception of average duration of fixations, significant changes in either reaction times or eye-movement patterns occurred only for recognized images. Moreover, the difference between recognized and not recognized images was not significant in the initial presentation, but it became evident in the delayed presentation. This would suggest that initially, all images receive similar “investment” in terms of processing time and visual attention, irrespective of whether they are subsequently
recognized or not. However, afterward, the trajectories of recognized and unrecognized images start to differ and the processing of recognized images starts to display signs of optimization, while no such changes occur for the unrecognized images. We found shorter reaction times, smaller dispersion and fewer fixations for recognized images in the delayed presentation, compared both to recognized images in the initial presentation and unrecognized images in the delayed presentation. The only exception was the duration of fixations – here we found a significant effect of presentation, but not recognition. In other words, the duration of fixations was higher in the delayed presentation compared to the initial presentation, irrespective of whether the images were recognized or not. This may indicate that the measure of fixation duration was not sensitive enough to reveal differences between recognized and unrecognized images.

However, we found that overall, recognized images, compared to unrecognized images, were processed more quickly and the related eye-movement patterns were more efficient and economical. These results are consistent with the proposed hypothesis that optimization of the representation takes place after recognition is achieved. This indicates that information reduction results from the acquisition of a criterion for selecting what is relevant at the moment. In other words, knowing what to look for means also knowing what to ignore. For this reason, we would expect that absence of recognition would mean the absence of the criterion that allows classifying information as relevant and irrelevant. As a result, the unrecognized images do not benefit from the “economies of experience”. It is also important to note that recognition should be seen as the trigger rather than the beneficiary of the optimization, given that optimization is contingent on recognition. In other words, recognition induces more efficient and sparse way of looking at the stimuli.

4.4 The temporal patterns of eye velocity

In order to further delineate the relationship between recognition and optimization of processing, we performed analyses of changes in the eye-movement velocity within the course of the image presentation. We found that each subsequent presentation was related to lower eye-movement velocity (Fig. 3e). Similarly, we found that recognized images were related to lower velocity, but only in the last 500 ms of the delayed image presentation (Fig. 6). The precise course of changes in eye-movement is depicted in Fig. 7. The patterns of eye-movement velocity in the first 500 ms (or 50% of eye samples) of image presentation were not sensitive to either experience or recognition, and the velocity is very low while increasing with time. According to Mannan et al. (1995), initial eye-movements are automatic in nature.
ECONOMIES OF EXPERIENCE

and determined by spatial characteristics of the image. However, in the second half of the delayed presentation, the difference between velocities for recognized and unrecognized images was significant, with unrecognized images related to higher velocities. The differences between recognized and unrecognized images increased with time and were at their largest at the offset of stimulus presentation (Fig. 7). This result is consistent with several studies reporting a decrease in saccade velocity in response to stimulus repetition (Chen-Harris, Joiner, Ethier, Zee, & Shadmehr, 2008; Straube, Fuchs, Usher, & Robinson, 1997) and to an increase in task difficulty (Galley, 1993).

Of particular interest is the time of the onset of eye deceleration in response to recognized and unrecognized images, defined as the moment when eye velocity starts decreasing (Fig. 7). For recognized images, the onset of eye deceleration is significantly earlier than for unrecognized images. This may indicate the slowing down of the search process for image-disambiguating cues once the image is interpreted.

4.5 Does “looking the right way” facilitate recognition?

Our third hypothesis was derived from the adaptive scanning hypothesis by Myers and Gray (2010). We predicted that the degree of similarity of eye-movement patterns between two presentations will be higher if they were related to the same behavioral response (either recognized or unrecognized). The results of our study confirm that cases, where the response was the same for each of the two presentations of the same image, were accompanied by a higher degree of eye-movement similarity, than cases of different responses. This occurred both between the two degraded presentations we compared (initial and delayed) and between the compared degraded vs. undegraded presentations. Additionally, eye-movement patterns accompanying the two degraded presentations were more similar to one another than to the undegraded presentation, even though they were further away in time. Finally, eye-movement patterns accompanying the undegraded presentation were more similar to the delayed presentation than to the initial presentation. That means that experience-induced perceptual learning leads to a change in eye-movement patterns as well as the change in recognition rates. In other words, having experience with the disambiguated image facilitates recognition of its degraded version but also changes the eye-movement patterns to ones more similar to the way of looking at the disambiguated image. However, our study does not prove that this change in eye-movement patterns facilitates recognition. It could be that the representation of the image is combined with a memory trace of eye-movements’ sequence, as suggested by Noton and Stark (1971). As a result, the top-down expectation consists of both scene/object
identity and the related motor sequence of eye-movements. The question is whether following that memorized pattern aids recognition. For example, Jarodzka, van Gog & Gerjets (2013) reported that superimposing experts’ eye-movement patterns on complex visual stimuli facilitated learning and improved performance of novices, which would suggest that “the right way of looking” does help perceptual learning.

4.6 Conclusion
The study confirmed that eye-movement dispersion decreases with each stimulus presentation. Additionally, we found other “economies of experience”- namely, a decrease in the number of fixations, an increase in the duration of fixations and a decrease in eye velocity. This effect was contingent on recognition (with the exception of fixation duration), indicating that the selection of information from the stimulus was informed by the identity of the presented object.

Additionally, once recognition rates started deteriorating from near-ceiling levels, the eye-movement measures continued to improve, in parallel to a decrease in reaction times. We concluded that perceptual learning leads to a decrease in the amount of information extracted from the environment, resulting in leaner processing and more compact eye movement patterns, supporting the information reduction hypothesis (Haider & Frensch, 1999a). The purpose of this reduction is to optimize processing even after recognition had been achieved, to decrease the cost and effort related to perception.

Furthermore, we reported temporal course of changes in eye velocity. We found that the eye velocity in the initial period of image presentation (first 500 ms) was insensitive to experience and recognition, and as such is likely automatic and contingent on spatial features of the stimulus. However, in the last period of stimulus presentation, there was a significant difference between recognized and unrecognized images in the delayed presentation. There was also a significant difference in the deceleration onset, which happened earlier in the case of recognized images.

Finally, our study demonstrates that after exposure to the undegraded stimulus eye-movement patterns accompanying the degraded and undegraded stimuli become more similar. Moreover, the similarity was higher if the behavioral responses were the same- i.e. if the image was either recognized or unrecognized after both presentations. In our view, this constitutes evidence that eye-movement patterns can be modified and adapted to optimal
processing of a given stimulus via experience-driven perceptual learning, supporting the adaptive scanning hypothesis (Myers & Gray, 2010).

We conclude that experience-induced changes in eye-movement patterns reflect “economies of experience”, similar to the priming - related formation of sparser neural representations and reduction in reaction times - i.e. optimization with the purpose of more efficient and less costly visual processing.

References


ECONOMIES OF EXPERIENCE

Society Convergence on Computer Vision and Pattern Recognition (CVPR’05)
Workshops. San Diego, CA.


ECONOMIES OF EXPERIENCE

Figure captions

Fig. 1 about here
Examples of stimuli used in the study

Fig. 2
Schematic diagram of a single session in the study.

Fig. 3
A. Recognition rates, b. reaction times, c. average number of fixations, d. the average duration of fixations, e. eye-movement velocity and f. eye-movement dispersion for all four presentations. Error bars denote within-subjects 95% confidence intervals (Loftus & Masson, 1994).

Fig. 4
Eye-movement dispersion in all four presentations for social and non-social stimuli examples (points represent eye-movement samples of all participants looking at a given image). The inner circle encloses 2/3 of all eye-movement samples and the outer circle encloses 90% of eye-movement samples.

Fig. 5
Differences in the four reported measures between “Yes” and “No” responses for the initial and delayed presentations. A. reaction times, b. eye-movement dispersion, c. average number of fixations, d. the average duration of fixations Error bars denote within-subjects 95% confidence intervals.

Fig. 6
Differences in eye movement velocity between “Yes” and “No” responses for a. the first and b. the last 500 ms of image presentation, in the initial presentation and delayed presentation. Error bars denote within-subjects 95% confidence intervals.

Fig. 7
The dynamics of eye velocity in the course of the image presentation (100% of eye samples is equivalent to 1000 ms).

Fig. 8
Levels of eye-movement patterns dissimilarity between initial and delayed presentations, between delayed and undegraded and between undegraded and initial, separately for cases when responses to both presentations were the same (“same responses”) and for cases when the responses for each presentation differed (“different responses”). Error bars denote SEM.
Figure 1

- Degraded
- Undegraded

- Carnival dancers
- Girl in the orchard
- Vase of flowers
- Toad
Figure 2
Figure 3

A. Example of a social stimulus - “clown”

Initial Presentation

Undegraded Presentation

Immediate-after-undegraded Presentation

Delayed Presentation

B. Example of a non-social stimulus - “butterfly”

Initial Presentation

Undegraded Presentation

Immediate-after-undegraded Presentation

Delayed Presentation
Figure 4

A. Recognition rates

B. Reaction times

C. Average number of fixations

D. Average duration of fixations

E. Eye-movement velocity

F. Eye-movement dispersion
Figure 5
Figure 6

A. First 500 ms of image presentation

B. Last 500 ms of image presentation

Eye Movement Velocity [deg/s]

Initial Presentation
   ---
   Not recognized images
   ---
   Recognized images

Delayed Presentation
Figure 7

Average eye-movement velocity [deg/s] vs. % Eye Samples

- a: onset of deceleration for recognized images
- b: onset of deceleration for unrecognized images

Legend:
- Dashed line: initial presentation, not recognized
- Solid line: initial presentation, recognized
- Dotted line: delayed presentation, not recognized
- Dash-dotted line: delayed presentation, recognized
Figure 8