# Effects of trial repetition in texture discrimination

Noa Ofen [b], Anan Moran [c], Dov Sagi [a,*]

[a] *Department of Neurobiology/Brain Research, Weizmann Institute of Science, Rehovot 76100, Israel*
[b] *Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[c] *Gonda Multidisciplinary Brain Research Center and Faculty of Life Sciences, Bar Ilan University, Ramat Gan, Israel*

## Abstract

Performance on the texture discrimination task improves with practice but was also shown to decrease between closely spaced sessions. Here we explored immediate changes in performance *within* a single session. We found that, after an initial increase, performance declined with further training within a single session. This deterioration in performance was smaller when the inter-trial interval was longer than 3 s. Performance recovered when targets were presented in new locations within the texture stimulus—thereby excluding a general fatigue process or adaptation to the stimulus light-intensity as an explanation for our findings. Further, the complete transfer of deterioration between eyes pointed to cortical origin. Deterioration was also found for task-irrelevant targets, indicating the involvement of a sensory mechanism. Collectively, these findings trace the deterioration of performance in the texture discrimination task, previously observed across several hours, to cortical events occurring during or immediately after stimulus presentation.
© 2007 Published by Elsevier Ltd.

*Keywords:* Adaptation; Perceptual learning; Spatial vision; Texture discrimination

## 1. Introduction

In general, repeated performance of a task leads to improved performance. This effect of practice has been shown in the texture discrimination task (Karni & Sagi, 1991; Karni & Sagi, 1993). Recently, however, it was found that practice can reduce performance in the texture discrimination task. Using the texture discrimination task, Mednick and colleagues (Mednick et al., 2002) showed that multiple training sessions within a single day led to decreased performance (Mednick et al., 2002). Their findings could not be explained by the effect of general fatigue, since sessions were spaced by several hours and performance recovered when the stimuli were switched to a new, untrained location in the visual field. Performance also changes within a single practice session. Within-session performance was reported to improve during the initial phase of learning, mainly in the first session (Karni

& Sagi, 1993), to be relatively stable during the following daily sessions (Karni & Sagi, 1993), or to decline during a second daily session (Mednick, Arman, & Boynton, 2005).

The aim of the present research was to examine within-session effects on performance. Of particular interest was the effect of different amounts of training within a single session on the performance within that session. Such effects might occur due to learning and sensory adaptation to the stimulus energy. In the absence of known mechanisms, a distinction between adaptation and learning can only be based on phenomenology; here we assume a demarcation based on task-relevancy. Sensory adaptation, such as contrast adaptation, is thought to be independent of the task performed and the adapting stimulus (Festman & Ahissar, 2004). Perceptual learning, however, depends on the task performed (Ahissar & Hochstein, 1993; Karni & Sagi, 1995) and affects not only the trained stimulus aspects but also the associated stimulus parts (Seitz & Watanabe, 2003).

Effects resulting from the amount of training within a session are not restricted to perceptual learning (Ofen-Noy, Dudai, & Karni, 2003). These effects may

---

have critical implications for constructing efficient training and for understanding the nature of neural experience-dependent modifications that accompany performance changes. The interaction of within-session effects and between-session effects is presently not known. A recent study with the texture discrimination task found a non-monotonic dependency of between-session improvement on the number of trials within a session (Censor, Karni, & Sagi, 2006). In particular, between-session improvement was not observed with sessions containing too many trials, pointing to a strong interaction between processes operating within and between sessions.

The present study was designed to evaluate the effects of increasing the number of trials on performance within a single session. To this end, we used a variation of a texture discrimination task (Karni & Sagi, 1991) in which participants had to determine whether two targets were aligned the same or differently (Fig. 1). We conducted four experiments, detailed below, that demonstrate instability in performance within a session. The observed instability suggests the existence of within-session adaptation processes in texture discrimination task.

## 2. Methods

### 2.1. Participants

All together, 26 individuals (age range 17–32, 7 males) participated in the experiments. Participants gave informed consent and were monetarily compensated for their time. The number of sessions in which data were collected for each of the experiments is detailed below. Depending on random assignment, each individual might have been tested multiple times in one or more of the experiments with the strict constraint that only one session was given in a single day.

### 2.2. Task and stimuli

We used a variation of a texture discrimination task (Karni & Sagi, 1991). Participants watched a computer screen from a distance of approx-



Fig. 1. Experimental trial. Trials were self-initiated and a fixation screen (a small '○' at the screen center, not shown) was followed by a stimulus screen (upper left) that included two targets composed of 3 45° line segments, embedded in an array of 19 × 23 horizontal segments. The stimulus was presented for 18 ms and was followed by a mask (array of 19 × 23 V-shaped segments, upper right) presented after a variable SOA (Stimuli-to-mask Onset Asynchrony).

imately 1.2 m in a dark room. Each trial was initiated by the participant pressing the middle button on a three-button mouse after a white fixation circle appeared in the middle of the screen. The stimulus was composed of a 19 × 23 array of horizontal white line segments on a black background with two targets defined by orientation difference. The targets were horizontally or vertically arranged sets of three diagonal line segments (Fig. 1). The two targets were presented at equal distances from the fixation point either on the horizontal mid-line or, in different experiments, in diagonally opposing quadrants of the screen (i.e. upper-left and lower-right or lower-left and upper-right). Stimuli were presented for 18 ms followed by a variable blank interval (stimulus-to-mask onset asynchrony, SOA) that was followed by the presentation of a patterned mask (150-ms duration). The mask was composed of 19 × 23 V-shaped elements with the elements' orientation randomized during each trial. Participants made a two-alternative, forced-choice decision between 'same' or 'different' (clicking on either the right or left buttons of a three-button mouse). Participants responded 'same' when the three segments comprising the two targets were aligned similarly (i.e. both horizontal or both vertical) or 'different' when they were aligned differently (i.e. one horizontal and the other vertical). Feedback (a beep sound) was given for incorrect answers. Trials were grouped into blocks of constant SOA. There were four possible target arrangements presented within a block (two targets, each having two possible orientations), one of which was selected at random during each trial. Blocks were terminated when the number of presentations of each of the four arrangements exceeded a pre-defined number (N/4). The number of trials per block (N) for each experiment is provided in the Section 3, with the actual number of trials being slightly larger. The percentage of correct responses was calculated per block for the given number of trials (N) with equal numbers of the four possible target arrangements. No additional fixation task was employed [in contrast to Karni and Sagi (1991)].

### 2.3. Measurements

Performance was measured by the rate of correct responses in a block. Threshold SOAs in ms were calculated by interpolating the SOA for which 80% probability of a correct response would be obtained. Comparisons between conditions were done using Repeated-Measure ANOVA (models specified in text). Further statistical tests were done using 2-tail *t*-tests corrected for multiple comparisons (Bonferroni) when performed for more than a single pair.

### 2.4. Experimental procedures

*Experiment 1*: Fifteen participants (age range 18–27, 4 males) were tested multiple times on different days, providing a total of 69 sessions. Before the experiment proper, an individual threshold was determined for each participant. This threshold was determined via descending SOAs (blocks of successive 20 trials with descending SOAs; 400, 300, 200, 160, 140, 120, 110, 100, 90, 80, 70, 60, and 50 ms). In this initial threshold estimation testing was terminated when the participant was performing near the chance level.

In the experiment proper, participants were given 1000 trials using a single, above-threshold SOA. For an above-threshold SOA, we chose one in which the performance in the initial test was better than 90% correct responses. Performance was analyzed in blocks of 40 trials. Participants were instructed not to take voluntary breaks longer than 30 s in between trials. All together, we determined performance in 69 sessions. Across these sessions we were using variable test parameters, as described below.

Three test parameters were manipulated. First, we used binocular or monocular viewing (49 and 20 sessions, respectively). Second, we had participants with different amounts of previous exposure to the task. Prior exposure was the result of the same participants being tested on previous days, ranging from 0 (for the first day of testing) to 13 prior daily sessions. The third parameter was the single above-threshold SOA. This ranged between slightly above and greatly above the participants' thresholds (difference between threshold and selected test SOA: 5–292; SOAs: 70–400; thresholds: 55–275; range in ms). Using these multiple alterations in
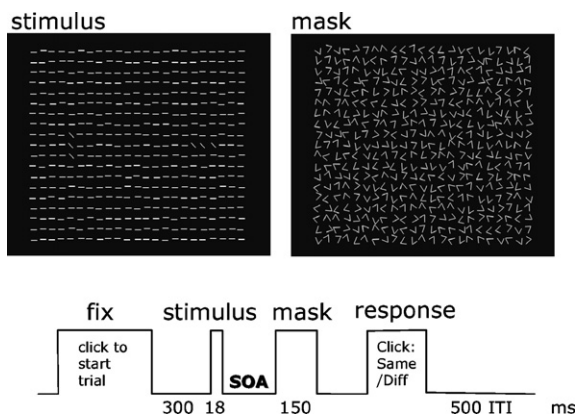
the test parameters allowed to assess the possible generality of any pattern of within-session performance arising due to repeated practice.

*Experiment 2*: Eight participants (age range 18–32, 1 male) were tested in a total of 13 sessions. Testing was given in blocks with descending SOAs in three phases (FEW1, MANY, FEW2; consisting of 20, 200, and 20 trials per block/SOA, respectively). The FEW conditions were administered as in the initial threshold estimation in Experiment 1. Only a subset of SOAs was used in the MANY condition, starting with an SOA for which above 90% correct responses was obtained. This was done to maintain overall session duration of about 1 h.

*Experiment 3*: Ten participants (age range 18–29, 4 males), in multiple sessions, were given an initial threshold estimation (as described above in Experiment 1) and then randomly assigned to one of 4 experimental conditions. Training with a single, above-threshold SOA was given (a) continuously, as described in Experiment 1 (CONS, $n = 11$), (b) with 1-min inter-block intervals (1MIN, $n = 16$), (c) with 5-min inter-block intervals (5MIN, $n = 22$), or (d) with a 3 s inter-trial interval (3SEC, $n = 15$). Training with a single SOA was then given for a total of 50 min. Training SOA was set at the lowest SOA for which performance was at, or above 90% of correct responses in a block of 20 trials given in the initial threshold estimation phase. Thus, experimental conditions CONS, 1MIN, 5MIN, and 3SEC consisted of 25, 17, 8, and 10 training blocks, respectively.

*Experiment 4*: Nine participants (age range 17–26, 5 males) were tested in the transfer tests and contributed to a total of 30 sessions. In these sessions, participants were randomly assigned to one of three transfer conditions: transfer across the eyes ($n = 13$); transfer across locations in the visual field ($n = 15$), or transfer across locations with task-irrelevant targets present throughout the testing ($n = 2$). Initial threshold estimations (administered as described above in Experiment 1) were made for both the 'before' and 'after' transfer stimuli conditions. In the eye condition the monocular viewing was enforced by a patch comfortably placed atop one of the eyes and targets appeared horizontally to the left and right of the fixation mark. In the location condition the targets were located in diagonally opposing quadrants. When task-relevancy was tested, targets were presented in four quadrants, each with independently chosen orientations, with target-pairs formed between diagonally opposed quadrants. Training with a single SOA was given for about 25 blocks followed by 3 blocks of transfer. The training SOA and the transfer SOA were set as in Experiment 3. The initial viewing eye and specific (or task-relevant) locations were counterbalanced across sessions. Three data points were calculated for each transfer test session: the mean performance (percent correct) in the first 3 blocks of practice (First); in the last 3 blocks before the transfer phase test (Last); and in the first 3 blocks of the transfer test (Switch).

# 3. Results

## 3.1. Experiment 1: Performance fluctuations with a single, above-threshold SOA

When testing perceptual thresholds, performance is often reduced experimentally. For example, many studies of texture learning estimated threshold by reducing the SOA (the so-called "descending" method) (Censor et al., 2006; Karni & Sagi, 1991; Mednick et al., 2002). However, mere repetition might lead to reduced performance by an adaptation-like process (Grill-Spector, Henson, & Martin, 2006), triggered by the repetition of the stimulus presentation and task. In Experiment 1, we tested the performance on many successive trials with a constant SOA. An initial phase of threshold estimation, a relatively rapid descending method (20 trials per SOA, see Section 2), was followed by a second phase in which a single, above-threshold SOA was used in 25 blocks of about 40 trials each. Remarkably, when a sin-

gle, above-threshold SOA was given, the participants were not able to attain a steady high level of performance across the 25 continuous blocks. Instead, initial near perfect performance in the first block (94% ± 5, percent correct responses ± SD) was followed by fluctuations between high and low levels of performance, resulting in, for the 25 blocks, an average of 76% ± 8 (range: 96–60% across sessions, 69 sessions). Marked deterioration from the initial level of performance was evident both at the individual session level (see examples of 11 different sessions, Fig. 2a) and in the averaged data (across all 69 sessions, Fig. 2b).

A consistent but small effect of improvement in performance is seen between the first and the second blocks (block 1: 94% ± 5, block 2: 96% ± 5, $p < .01$), possibly reflecting within-session learning often observed in the first days of training (Karni & Sagi, 1993). Consistent with this notion, the magnitude of these improvements was correlated with the day of testing, such that it was larger in the initial days of testing and decreased in sessions conducted on later days ($r = .29$, $p < .05$). However, not surprisingly, in the present study this improvement is minimal because the initial performance level is close to perfect, making it hard to observe improvements (ceiling effect). The main effect of interest here is the robust deterioration in performance during later stages of within-session practice.

An examination of the data from individual sessions showed an initial deterioration that was often followed by fluctuations between high and low levels of performance. These transient improvements that followed the first decline in performance are barely discernable in the group-averaged data, indicating that there was no consistent fluctuation cycle across participants. The first drop in performance was relatively constant across sessions, occurring typically after about 200 trials (4–5 blocks). This first drop remained evident in the averaged data (block 4: 91% ± 8%, $p < .05$; block 5: 86% ± 12%, $p < .001$; $p$ values of 2-tail *t*-test against performance in block 1, corrected for multiple comparisons). The fluctuations in performance cannot be accounted for by the notion of noisy measurement due to a limited sampling size, since the standard deviation for each individual session was much higher than the expected standard deviations for an individual session (assuming binomial distribution, calculated as $(p(1-p)/N)^{0.5}$, with $p$ being the estimated mean probability of a correct response in a session and $N$ being the number of trials in a block of trials; dashed line in Fig. 2c).

Finally, we would like to point to the generality of the observed phenomenon. The 69 sessions taken into account in this experiment were administered with a wide range of test parameters (see Section 2.4). These sessions were conducted with binocular or monocular viewing, across different amounts of previous exposure to the task, and with a large range of SOAs (70–400 ms). All these conditions yielded similar effects on performance (e.g., compare the 100-ms and 400-ms conditions in Fig. 2b). Importantly, these test parameters did not influence the performance on the first block, since the test SOA was always set to be
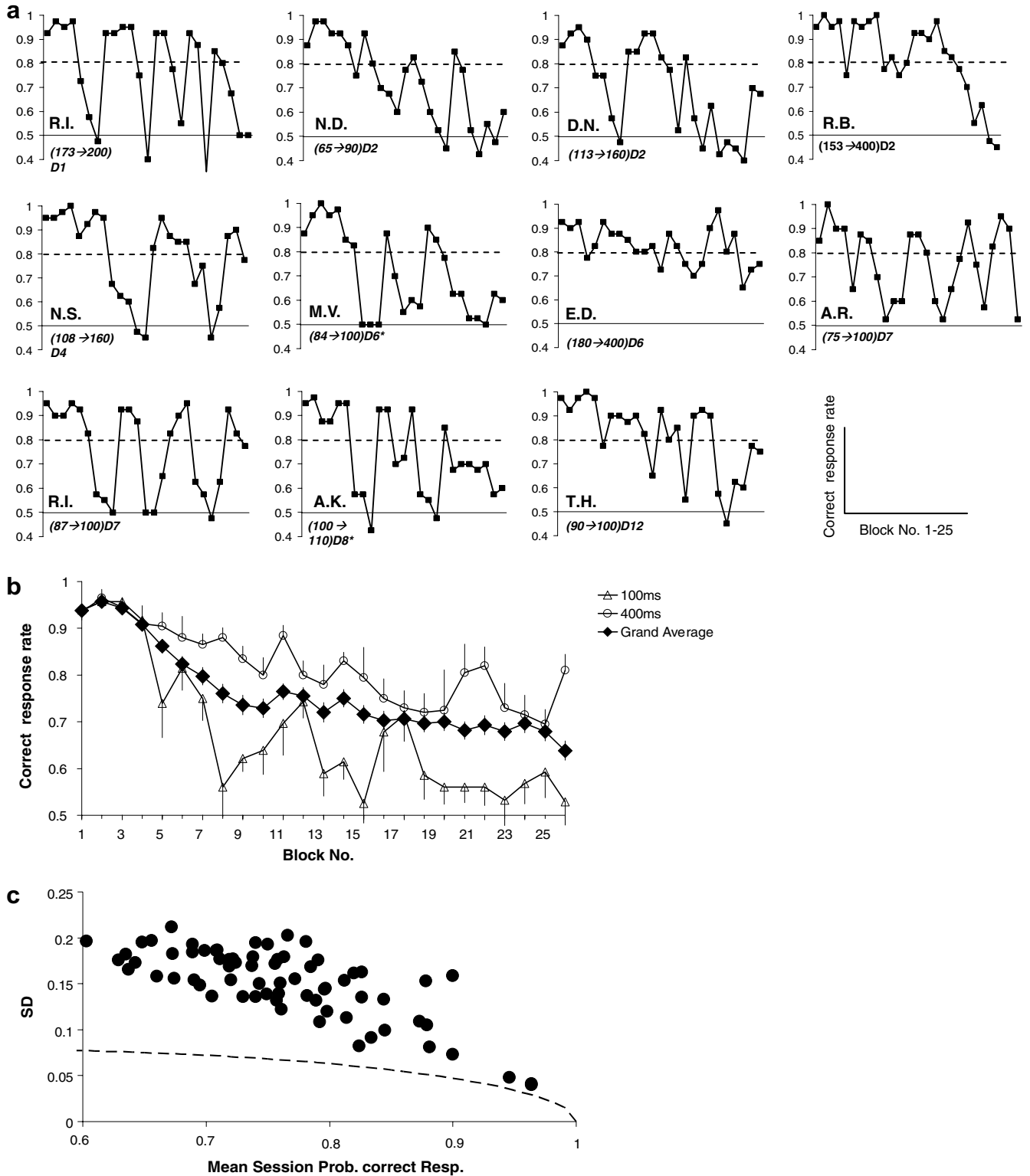
Fig. 2. Performance across 25 blocks as tested in Experiment 1 with a single, above-threshold SOA. (a) Examples of correct response rates (fractions) across the 25 blocks (11 representative examples out of a total number of 69 sessions) given to different participants, with variable perceptual thresholds and with variable amounts of previous exposure to the task (on previous days), using different SOAs and under either monocular or binocular vision. Inside brackets: *Threshold*, in ms, as estimated in an initial phase prior to the 25-block test; *test SOA*, in ms, the SOAs chosen for the 25-block test. *D#* denotes the number of previous daily sessions plus the day in which the session was performed; * denotes monocular viewing. The dashed line in 0.8 represents the defined threshold. (b) Mean performance for the 25 blocks across all sessions (Grand Average, filled diamonds) and for sessions in which subjects were tested with either a 100-ms (open triangles, $n = 7$) or 400-ms (open circles, $n = 5$) SOA on days 3–6 (to equate skill level). Error bars are SE across individual sessions. (c) The fluctuations are not due to measurement noise, since the standard deviation for the individual sessions greatly exceeds the expected standard deviations, assuming that the noise was attributed to a limited number of trials in a block (dashed line).

above the participant's threshold, as assessed in the initial phase of the session. The overall performance across the 25 blocks was only influenced by the difference between the participant's threshold and the test SOA (ΔSOA, 5–292, the range in ms). The averaged percent of correct responses increased with ΔSOA ($r = .48$, $p < .001$), and the variance decreased with ΔSOA ($r = -.52$, $p < .001$). The number of daily sessions prior to the test did not correlate with the average or the variance in performance across sessions ($ps > .13$). In addition, there were no differences in performance between monocular and binocular viewing once controlled for ΔSOA ($ps > .21$).

### 3.2. Experiment 2: Threshold estimation

In Experiment 1, we found that performance within a session was not stable and eventually deteriorated. This might have an important 'side-effect' regarding the estimation of thresholds. Typically, threshold is estimated by successive blocks of descending SOAs. At each SOA level, a certain number of trials are given. The results of Experiment 1 strongly suggest that the number of trials given per SOA might critically affect the measured threshold. In Experiment 2 participants were given 3 tests of thresholds with either a few trials (20: FEW1, FEW2, given at the beginning and end of the session) or many trials (200: MANY given in-between) per SOA in successive blocks with descending SOAs ($n = 13$). Overall, as shown in Fig. 3, the participants obtained lower thresholds under the FEW test conditions compared with the MANY test condition (FEW1: $87.9 \pm 4.6$, MANY: $102.7 \pm 3.7$, FEW2: $95.6 \pm 4.3$; mean threshold SOA in ms $\pm$ SD). Repeated measures analysis of variance (Condition: FEW1, MANY, and FEW2, threshold SOA in ms as a dependent measure) confirmed the significant effect for Condition ($F_{(2,24)} = 25.5$, $p < .001$). Further pair-wise comparisons showed significant differences under all conditions. Lower thresholds were obtained when testing with only 20 trials per SOA compared to when testing with 200 trials per SOA (MANY vs. FEW1: $p < .001$; MANY vs. FEW2: $p = .005$). This result is consistent with previous findings concerning the relationship between performance (measured threshold) and the number of trials that constitute the measurement (Censor et al., 2006).

In addition, we found that thresholds assessed with 20 trials per SOA were higher at the end of the session compared to the beginning of the session (FEW1 vs. FEW2: $p = .02$). The increase in threshold at the end of the session may indicate long-lasting within-session deterioration component that was independent of the number of repeated trials given while assessing the performance threshold. Thus, performance is affected both by the number of trials that constitutes a specific measure and the number of preceding trials in the session's 'history'.

### 3.3. Experiment 3: Spacing trials in time

In Experiment 1 performance deteriorated within a session when the same stimuli were given continuously, with an average inter-trial interval of about 2 s. Experiment 3 was aimed at assessing the occurrence of within-session deterioration in performance with varying presentation rates. Specifically, we tested whether the within-session drop in performance is selectively due to cumulating trials consecutively (CONS, with no time intervals in between) or whether the deterioration would occur even if within-session practice is spaced in time. We achieved this manipulation by maintaining the overall session time fixed (at 50 min) and inserting time intervals either between blocks (1 min, 1MIN or 5 min, 5MIN) or between trials (3 s, 3SEC). Differences in performance between continuous testing and spaced testing would determine the role of a general fatigue mechanism in accounting for the deterioration in performance, since, under all the experimental conditions, training is given for an overall fixed amount of time (50 min).

The results of experiment 3 showed that inter-block intervals of 1- and 5-min and 3-s inter-trial intervals both lead to less performance deterioration, when compared to the continuous condition (Fig. 4). Statistical effects were calculated in two ways, equating for time in sessions and equating for the number of task repetitions. First, for each individual, the 50-min session was subdivided into 5 time bins, each representing approximately 10 min. A $4 \times 5$ Repeated Measures model was used (Group: CONS, 1MIN, 5MIN, 3SEC; Time: 10-min time bins; percentage of correct responses as a dependent measure); it showed both Group ($F_{(3,65)} = 11.01$; $p < .001$) and Time ($F_{(4,260)} = 70.57$; $p < .001$) main effects and a significant Group × Time interaction ($F_{(4,260)} = 6.14$; $p < .001$). Further testing for the Group differences within each time bin showed a significant effect for the Group in all but the first 10-min time bins. Comparing between groups within each of the remaining time bins showed that the 5MIN group outperformed the CONS group in all four time bins ($p < .001$), whereas the 1MIN group showed less deterioration compared with the continuous group only in the 10–20-min time bin ($p = .03$). The 3SEC group outperformed the CONS group in both the 10–20-min and 40–50-min time bins ($p < .001$, $p = .02$, respectively). These results indicate overall less deterioration for both 1-min and 5-min
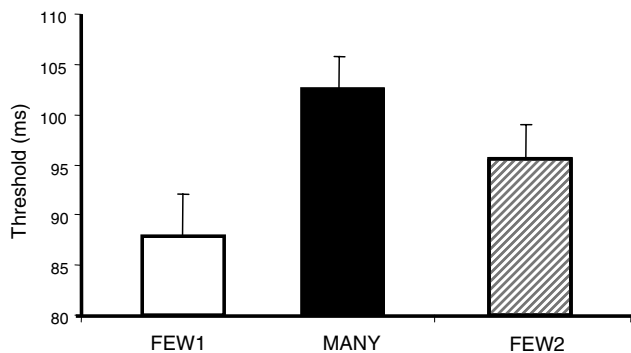


Fig. 3. Mean group threshold SOAs (ms, error bars represent SE across individual tests) for testing with 20 trials per SOA (FEW1, open column; FEW2, striped column) or with 200 trials per SOA (MANY filled column).
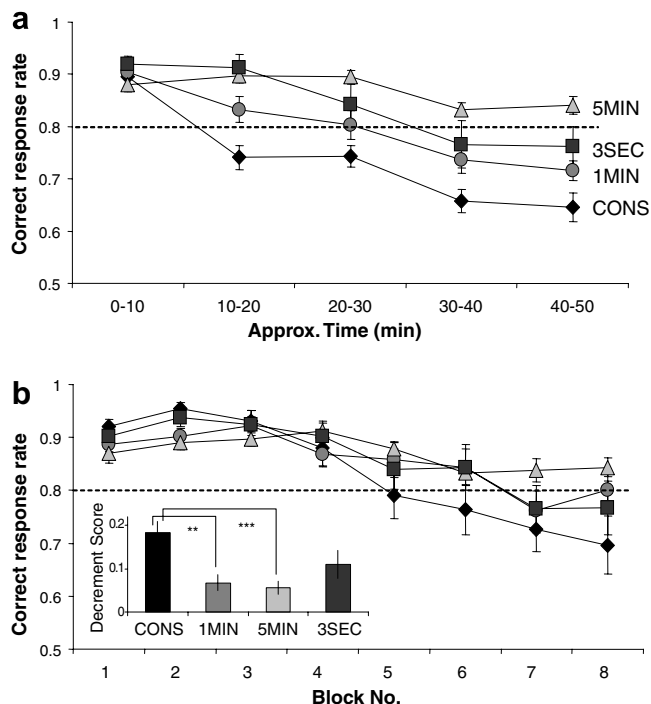
Fig. 4. Within-session performance: continuous vs. time-spaced testing. Group mean correct response rate plotted for five consecutive time bins of approximately 10 min each (a), or for comparable block numbers (b). CONS, consecutive condition—black diamonds, 1MIN, 1-min inter-block interval — gray circles; 5MIN, 5-min inter-block intervals—gray rectangular; 3SEC, 3-s inter-trial intervals—dark gray squares. Error bars represent SE across individual sessions (***, $p < .001$;**, $p < .005$).

inter-block intervals as well as for 3-s inter-trial intervals compared with continuous training.

Next, we tested for Group differences given a fixed number of blocks (Fig. 4b). Only the first 8 blocks were considered, to match the maximum number of blocks in the condition most spaced in time (5-min inter-block interval). $4 \times 8$ Repeated Measures model (Group as above, Block: blocks 1–8) showed no Group effect but a significant Group × Block interaction ($F_{(21,462)} = 2.43$; $p = .003$) that may reflect similar levels of performance across groups in the initial few blocks and different levels of deterioration across groups thereafter.

In order to further characterize this interaction, we computed a decrement score as the average correct response rate in the first 4 blocks minus the average correct response rate in the last 4 blocks. This decrement score was smaller in groups 1MIN and 5MIN compared with the CONS group ($p < .005$; $p < .001$, respectively). The decrement score of group 3SEC did not differ from that of the continuous group. Thus, introducing 3 s between trials reduced the overall degree of within-session deterioration only when differences were assessed across the duration of the whole session. Group 3SEC did not differ from group CONS if the same number of task repetitions were compared, suggesting that the deterioration is maintained for up to 3-s inter-trial intervals. Inter-block intervals reduced deterioration both for respective time frames and similar amounts of

task repetition. Overall, the results of Experiment 3 showed that the time afforded between blocks of trials allowed for significant recovery, thereby attenuating performance decrements. The obtained temporal parameters, in addition to being useful descriptors of the dynamics of the processes involved, have practical implications since parameters such as inter-trial and inter-block durations are rarely controlled in experiments. Interestingly, some neurophysiological studies use a 3-s minimal inter-trial interval to avoid adaptation in single-cell recordings from the visual cortex (Barlow, Kaushal, Hawken, & Parker, 1987).

### 3.4. Experiment 4: Deterioration specificity—viewing eye, spatial location, and task-relevancy

The specificity of the observed within-session deterioration was tested using three transfer conditions: across eyes (inter-ocular); across locations in the visual field; and across locations in the visual field with the addition of irrelevant targets present during practice. Similar levels of low performance before and after the transfer manipulation implies that the deterioration transfers across the measured dimension. Such a finding would indicate that the switched dimension is not a critical contributor to the deterioration. Better performance following a specific transfer manipulation would indicate that the manipulated dimension was critical for the manifestation of the deterioration. To determine transfer we used the mean performance in the last blocks before the transfer (Last) and the first blocks of the transfer test (Switch) (Fig. 5). The deterioration in performance, observed during the session, was transferred across eyes (Last: 59% ± 10%; Switch: 64% ± 11%; probability of correct response ±SD; $p = .31$). In contrast, the deterioration did not transfer across locations in the visual field (Last: 57% ± 12%, Switch: 93% ± 4%; $p < .001$). Finally, when task-irrelevant targets were present during the session, the deterioration did transfer to the new task-relevant locations following the switch (Last: 55%, 68%; Switch: 58%, 58%, subject RI and DN, respectively).
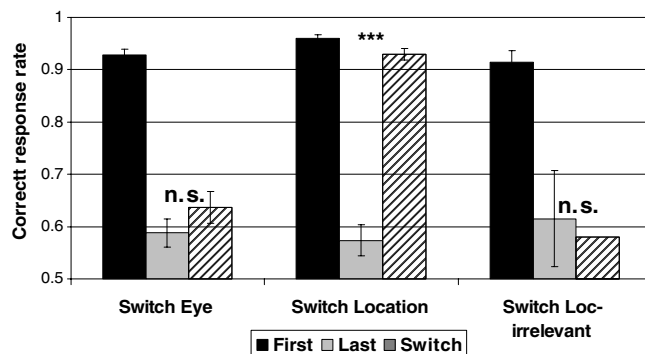


Fig. 5. Group mean correct response rates in the first 3 blocks (First- black columns), the last 3 blocks before the transfer test (Last-gray columns) and the first 3 blocks of the transfer test (Switch-striped columns) for the eye transfer condition, the location transfer condition, and the location with task-irrelevant stimuli present throughout the testing transfer conditions. Error bars represent SE across individual sessions (***, $p < .001$).

## 4. Discussion

In the experiments described here, changes in performance in a texture discrimination task were tested within a practice session, keeping stimulus parameters (SOAs) constant. Previous studies using a limited number of trials (<200) with fixed SOAs showed improved performance during the first practice session and stable performance in later sessions (Karni & Sagi, 1993). This work, however, presents evidence that within-session performance is highly unstable. With increased amounts of training on the same task, participants' performance showed the following pattern: a slight increase in performance during the first 100 trials, followed by marked fluctuations, with a continuous deterioration in performance.

The high level of initial performance allowed us to expose the large deterioration in performance observed during the practice session. Such deterioration could be attributed to several factors involved in the task, ranging from low-level light and contrast adaptation, to high-level limitations related to participants' alertness. We used the feature and task selectivity of the performance deterioration to rule out some known processes.

We showed that within-session deterioration was not restricted to the viewing eye, and therefore could not reflect retinal adaptation but rather a cortical process. Was adaptation to occur in the retina, than we would expect to find that switching eyes would remove the deterioration. It did not. The specificity of the deterioration to the visual field location suggests the deterioration process involved visual cortical regions where a retinotopic mapping of the visual world is maintained. However, this adaptation is driven by stimulus presentation regardless of the task-relevancy of the stimuli, which diminish an intentional attentional top-down explanation. This suggests that the phenomenon depends on sensory sensitivity that is not driven by specific task demands, possibly the result of a low-level cortical process.

Adaptation processes with repeated stimuli presentations have been shown to affect performance on psychophysical tasks. For example, psychophysical studies have demonstrated that extended exposure to high-contrast stimuli elevates the detection threshold of succeeding stimuli, a collection of effects studied under the title of "adaptation." Our results are in agreement with the time scale of contrast adaptation Contrast adaptation effects with short duration adaptors are largely absent 2 s after the adaptation period (Foley & Boynton, 1993), whereas here, with the brief stimuli used, the performance decrements were reduced with 3-s inter-trial intervals.

In the present study, the target duration was 18 ms and was followed by a mask of 150 ms duration. The mask was constructed with randomly oriented line elements presented in all stimulus locations; thus adaptation caused by the mask cannot explain the location specificity observed here. Retinal-light adaptation is ruled out in the current study, since the present effect was transferred between the eyes.

Contrast adaptation is known to have a cortical component; however, with the high-contrast stimuli used here, its effect is reduced and is broadband in orientation (Snowden & Hammett, 1992). On the other hand, a recent study showed that adaptation affects the contrast detection threshold of a target embedded in external noise (note that our stimulus also includes external noise—the mask), with the threshold elevation being constant across a large range of external noise contrasts (Dao, Lu, & Dosher, 2006). Dao et al. (2006) found a sharp orientation tuning for the adaptation effect. Orientation selectivity is of particular relevance to the phenomena described here because performance deterioration seems to be specific to orientation, showing absolutely no transfer across an orientation difference of 45°, since locations that were extensively stimulated by horizontal background lines do not show any deterioration effects when tested with the diagonal target-elements in the location transfer test. However, it is possible that, in the present experiments, the adaptation affects second-order processes acting on texture boundaries (Sagi, 1990). Such adaptation effects were recently observed in the human visual cortex (Larsson, Landy, & Heeger, 2006), with attention distracted from the stimulus. Consistently with these results, we observed performance deterioration for targets repeatedly presented without a task.

Maintaining focused attention over extended periods of time has been studied in the context of vigilance attention (Ballard, 1996; Parasuraman, Warm, & See, 1998). In these accounts, vigilance was defined as a state of readiness to detect and respond to certain small changes occurring in the environment at random time-intervals. Vigilance was often assessed in tasks requiring the detection of transient, infrequent, unpredictable events over long periods of time. In situations such as these, the quality of attention was fragile and declines over time, an outcome known as the vigilance decrement. However, vigilance decrement might be only loosely related to the fluctuation in performance described here, since our displays were frequent and predictable. In fact, each trial was initiated by the participant, so that the participant was prepared as much as possible for the upcoming display of stimuli. Furthermore, vigil decrements are typically presented as summarized performance across a testing session. A single measure for the whole session (or large chunks within a session) will not capture the dynamic nature of the phenomena characterized in the current study. Our findings clearly show that performance did not simply become gradually worse; rather, we observed marked fluctuations in performance, as indicated by the large variability scores. Due to the predictability of the stimuli presentation in time and the dynamic nature of performance fluctuations, we believe that vigilance does not account for our findings. The specificity of the deterioration regarding location in the visual field is taken as further evidence against an account of general fatigue. The robust recovery following the location switch argues for a dynamic change in sensitivity within sessions that might reflect repetition-dependent processes in low-level visual cortical areas.

The link between within-session performance fluctuations detailed here and between-session practice-related improvements is yet to be tested. Fluctuations in performance within a session might have critical implications regarding changes in performance across sessions (i.e. learning and consolidation processes). A recent study showed that between-session improvements in performance critically depend on the number of trials within a session, with improvements absent when the number of trials was large (Censor et al., 2006). There are indications that such a failure to improve performance can be restored by sleep (Censor et al., 2006). Deterioration in performance on the texture task, between sessions and within a day, was found to be abolished by taking a nap (Mednick, Nakayama, & Stickgold, 2003; Mednick et al., 2005). Furthermore, across-night improvement was found to depend on the integrity of night sleep (Karni, Tanne, Rubenstein, Askenasy, & Sagi, 1994; Stickgold, James, & Hobson, 2000). A recent study showed a link between local slow-wave activity during sleep and improved task performance the following day (Huber, Felice Ghilardi, Massimini, & Tononi, 2004). It is possible that such local slow-wave activity during sleep is critical for diminishing the cause of the deterioration occurring during continuous within-session task repetitions.

In another study, within-session decrements in performance on a perceptual acuity task were shown to be accompanied by lower levels of synchronous activation of cortical neurons (Ludwig & Skrandies, 2002). These within-session changes were, in turn, associated with between-session effects if the performance in the following session improved (Ludwig & Skrandies, 2002).

Brain imaging studies typically find that neuronal responses to the second of a pair of briefly presented visual stimuli are smaller than would be expected from the response to a single stimulus alone (Boynton & Finney, 2003; Grill-Spector et al., 2006; Grill-Spector & Malach, 2001). These reductions in brain activations were thought to reflect adaptive tuning processes that enhance efficient processing of the stimuli (Wiggs & Martin, 1998). However, the relationship between these neuronal processes and task performance is not clear. Although our experimental designs (Experiments 1, 3, and 4) were not optimized to show fast, within-session improvements, we have noted improvement in performance from the first to the second block (directly tested in Experiment 1). These improvements were greater during the initial days of testing, consistent with previous findings. Further study is needed to determine whether the initial improvement and the ensuing deterioration share common underlying mechanisms.

When discussing the similarities between the within-session decrements and perceptual learning, one may note the lack of eye specificity in our data. This result is in agreement with previous findings showing between-session deterioration (Mednick et al., 2005). Performance gains obtained within a session in a texture discrimination task transferred between eyes, whereas gains obtained between sessions were eye-specific (Karni & Sagi, 1991; Karni & Sagi, 1993; Walker, Stickgold, Jolesz, & Yoo, 2005). Although this behavior might vary between tasks (Ahissar & Hochstein, 1996; Ahissar & Hochstein, 1997; Fahle, 2004; Fahle, Edelman, & Poggio, 1995; Karni & Sagi, 1991) and testing conditions (Ahissar & Hochstein, 1996; Ahissar & Hochstein, 1997), our findings raise the possibility that perceptual deterioration involves mechanisms similar to those resulting in fast, within-session gains. The location selectivity we observed here is a property of both within-session and between-sessions improvements observed with a variety of visual tasks (Ahissar & Hochstein, 1996; Fahle, 2004; Fahle et al., 1995; Karni & Sagi, 1991; Karni & Sagi, 1993), although, again, the details depend on the task and methods used (Ahissar & Hochstein, 1996; Ahissar & Hochstein, 1997).

Finally, task-relevancy was shown to be critical for texture learning (Karni & Sagi, 1993), but was not necessary for the development of within-session adaptation observed here. Note that the task-relevant targets in the present experiments were not correlated with the simultaneously presented task-irrelevant targets. Thus, associations could not have been developed between the different targets, as observed by Seitz and Watanabe (2003). Hence, it is unlikely that our findings represent a global reinforcement process that affects both task relevant and irrelevant targets (Watanabe, Nanez, & Sasaki, 2001). Overall, the transfer profile of the present deterioration effect suggests some similarity but also dissociation from effects that are associated with learning across sessions. Thus, although performance deterioration was observed under testing conditions very similar to conditions that produce learning, the way these two phenomena interact is yet to be determined.

Taken together, this study revealed the critical role played by the test parameters. The number of task repetitions and their temporal interval exert important effects on within-session performance. Thus, performance thresholds were lower when assessed with fewer trials, and participants could not maintain an above-threshold level of performance with additional task repetitions. The number of within-session task repetitions interacted with the timing of successive trial presentations. There was less deterioration when there were substantial time intervals between blocks or trials within the session. Moreover, the number of within-session task repetitions interacted with the specific spatial layout of the stimuli, as evident by the fact that switching the target locations in the visual field was shown to completely diminish within-session deterioration. Overall this study provides strong evidence that performance varies within-session, including the counterintuitive result that practice can worsen performance. These changes in performance during a learning session have been largely overlooked in the past. We stress the need to further study the effect of such performance variability on practice both within and across sessions.

## Acknowledgments

We thank Jeffrey Ellenbogen and Yael Adini for helpful comments on early versions of the manuscript, and Yoram Bonneh for the psy software. This work was supported by the Basic Research Foundation administered by the Israel Academy of Sciences and Humanities and by the Benoziyo center for Neuroscience at the Weizmann Institute of Science.

## References

Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *PNAS, 90*, 5718–5722.

Ahissar, M., & Hochstein, S. (1996). Learning pop-out detection: specificities to stimulus characteristics. *Vision Research, 36*, 3487–3500.

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387*, 401–406.

Ballard, J. C. (1996). Computerized assessment of sustained attention: a review of factors affecting vigilance performance. *Journal of Clinical and Experimental Neuropsychology, 18*, 843–863.

Barlow, H. B., Kaushal, T. P., Hawken, M., & Parker, A. J. (1987). Human contrast discrimination and the threshold of cortical neurons. *Journal of the Optical Society of America A, 4*, 2366–2371.

Boynton, G. M., & Finney, E. M. (2003). Orientation-specific adaptation in human visual cortex. *Journal of Neuroscience, 23*, 8781–8787.

Censor, N., Karni, A., & Sagi, D. (2006). A link between perceptual learning, adaptation and sleep. *Vision Research, 46*, 4071–4074.

Dao, D. Y., Lu, Z. L., & Dosher, B. A. (2006). Adaptation to sine-wave gratings selectively reduces the contrast gain of the adapted stimuli. *Journal of Vision, 6*, 739–759.

Fahle, M. (2004). Perceptual learning: a case for early selection. *Journal of Vision, 4*, 879–890.

Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research, 35*, 3003–3013.

Festman, Y., & Ahissar, M. (2004). Attentional states and the degree of visual adaptation to gratings. *Neural Networks, 17*, 849–860.

Foley, J. M., & Boynton, G. M. (1993). Forward pattern masking and adaptation: effects of duration, interstimulus interval, contrast, and spatial and temporal frequency. *Vision Research, 33*, 959–980.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*, 14–23.

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica (Amst), 107*, 293–321.

Huber, R., Felice Ghilardi, M., Massimini, M., & Tononi, G. (2004). Local sleep and learning. *Nature, 430*, 78–81.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *PNAS, 88*, 4966–4970.

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature, 365*, 250–252.

Karni, A., & Sagi, D. (1995). A memory system in the adult visual cortex. In B. Julesz & I. Kovács (Eds.), *Maturational windows and adult cortical plasticity, SFI studies in the sciences of complexity* (pp. 1–15). Reading MA: Addison-Wesley.

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J. J., & Sagi, D. (1994). Dependence on REM sleep of overnight improvement of a perceptual skill. *Science, 265*, 679–682.

Larsson, J., Landy, M. S., & Heeger, D. J. (2006). Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *Journal of Neurophysiology, 95*, 862–881.

Ludwig, I., & Skrandies, W. (2002). Human perceptual learning in the peripheral visual field: sensory thresholds and neurophysiological correlates. *Biological Psychology, 59*, 187–206.

Mednick, S. C., Arman, A. C., & Boynton, G. M. (2005). The time course and specificity of perceptual deterioration. *PNAS, 102*, 3881–3885.

Mednick, S. C., Nakayama, K., Cantero, J. L., Atienza, M., Levin, A. A., Pathak, N., et al. (2002). The restorative effect of naps on perceptual deterioration. *Nature Neuroscience, 5*, 677–681.

Mednick, S., Nakayama, K., & Stickgold, R. (2003). Sleep-dependent learning: a nap is as good as a night. *Nature Neuroscience, 6*, 697–698.

Ofen-Noy, N., Dudai, Y., & Karni, A. (2003). Skill learning in mirror reading: how repetition determines acquisition. *Cognitive Brain Research, 17*, 507–521.

Parasuraman, R., Warm, J. S., & See, J. (1998). Brain systems for vigilance. In R. Parasuraman (Ed.), *The attentive brain*. MIT Press.

Sagi, D. (1990). Detection of an orientation singularity in gabor textures: effect of signal density and spatial-frequency. *Vision Research, 30*, 1377–1388.

Seitz, A. R., & Watanabe, T. (2003). Psychophysics: is subliminal learning really passive? *Nature, 422*, 36.

Snowden, R. J., & Hammett, S. T. (1992). Subtractive and divisive adaptation in the human visual system. *Nature, 355*, 248–250.

Stickgold, R., James, L., & Hobson, J. A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience, 3*, 1237–1238.

Walker, M. P., Stickgold, R., Jolesz, F. A., & Yoo, S.-S. (2005). The functional anatomy of sleep-dependent visual skill learning. *Cerebral Cortex, 15*, 1666–1675.

Watanabe, T., Nanez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature, 413*, 844–848.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinions in Neurobiology, 8*, 227–233.