

Jun-Yun Zhang¹, Shu-Guang Kuai¹, Lu-Qi Xiao¹, Stanley A. Klein^{2,3}, Dennis M. Levi^{2,3}, Cong Yu^{1*}

¹ Department of Psychology, Beijing Normal University, Beijing, China, ² Department of Psychology, University of California, Berkeley, California, United States of America, ³ Department of Psychology, University of California, San Diego, California, United States of America

Perceptual learning of visual features occurs when multiple stimuli are presented in a fixed sequence (temporal patterning), but not when they are presented in random order (roving). This points to the need for proper stimulus coding in order for learning of multiple stimuli to occur. We examined the stimulus coding rules for learning with multiple stimuli. Our results demonstrate that: (1) stimulus rhythm is necessary for temporal patterning to take effect during practice; (2) learning consolidation is subject to disruption by roving up to 4 h after each practice session; (3) importantly, after completion of temporal-patterned learning, performance is undisrupted by extended roving training; (4) roving is ineffective if each stimulus is presented for five or more consecutive trials; and (5) roving is also ineffective if each stimulus has a distinct identity. We propose that for multi-stimulus learning to occur, the brain needs to conceptually “tag” each stimulus, in order to switch attention to the appropriate perceptual template. Stimulus temporal patterning assists in tagging stimuli and switching attention through its rhythmic stimulus sequence.

J. Y. Zhang, S. G. Kuai, L. Q. Xiao, S. A. Klein, D. M. Levi, C. Yu (2008) *PLoS Biol* 6(8): 197–210. doi:10.1371/journal.pbio.0060197

Introduction

Practice improves discrimination of fine visual features, such as contrast, orientation, vernier offset (e.g., the alignment of two lines), texture, etc. [1–6]. This process is referred to as perceptual learning and has been studied intensively in recent years because of its close links to neural plasticity [7,8]. As in other forms of learning, stimulus information needs first to be encoded and consolidated into memory and later to be retrieved in order for perceptual learning to occur. Indeed, much has been done to understand stimulus coding effects on memory consolidation in a number of relevant domains. It is now clear that there are separate phases during which memory traces are susceptible to perturbation, during which they stabilize, and during which they are enhanced [9–11]. However, little is known about the coding stages of visual perceptual learning.

In a typical perceptual learning study, the observer practices a discrimination task at a single stimulus level (e.g., 30% contrast for a contrast discrimination task or 45° orientation for an orientation discrimination task), and after a few sessions of practice, discrimination at this stimulus level is usually improved. However, if the observer has to simultaneously learn discrimination of multiple stimuli (e.g., four different contrasts or orientations), perceptual learning is disabled if the stimuli are presented in a random temporal order (roving) [3,6,12,13]. In contrast, if the same stimuli are presented in a fixed temporal pattern (temporal patterning), substantial learning takes place [3].

The contrasting effects of temporal patterning and roving point to the need for proper stimulus coding to enable multiple stimulus learning. In this study we compare the effects of temporal patterning and roving and their interactions at different stages of perceptual learning, in order to reveal some basic stimulus coding principles for perceptual learning. Specifically, we investigated the roles of stimulus rhythm in perceptual learning, the effects of interruption by roving on temporal-patterned perceptual learning during consolidation and retrieval, and the minimal number of consecutive trials of the same stimulus required to escape disruption of learning through roving. Understanding the

stimulus coding rules has broad implications since such multi-stimulus learning is often encountered in natural learning and is not limited to vision. Consider, for example, a baseball batter. In order to succeed, the batter needs to quickly learn to identify whether the pitch is likely to be fast or slow and whether it will curve or spin. To the best of our knowledge comparable research has not been done in other modalities, such as motor and auditory learning.

The contrasting effects of temporal patterning and roving also pose serious challenges to existing models of perceptual learning. The models based on the activities of primary visual cortex (V1) neurons, such as the model of Adini, Tsodyks, and Sagi [14,15], assume that training modifies the recurrent connections in V1 neurons, so that the local network becomes more sensitive to the trained stimulus. However, such V1-based models cannot easily explain the effects of roving and temporal patterning on multi-stimulus learning. This is because different stimuli are responded to by different sets of stimulus-tuned V1 neurons, and there is no reason to believe that training induced modification in independent local recurrent networks would be differently affected by stimulus roving and temporal patterning. Alternatively, response reweighting models, such as the Lu and Dosher model [4,16], suggest that practice improves the readout of the most relevant V1 neuron responses to the stimulus by reweighting the responses of various neurons. The response reweighting models are designed to explain single stimulus learning. In their current form they are not capable of

Academic Editor: *Journal of Experimental Psychology*, B

Received November 29, 2007; Accepted March 3, 2008; Published August 12, 2008

Copyright: © 2008 *Journal of Experimental Psychology*, B
 DOI: 10.1371/journal.pbio.0060197

Abbreviations: 2A, *Journal of Experimental Psychology*; B, *Journal of Experimental Psychology*

* *Journal of Experimental Psychology*: congyu@bnu.edu.cn



Author Summary

explaining the roving and temporal patterning effects on multi-stimulus learning.

A more relevant model to our multi-stimulus learning results would be Ahissar and Hochstein's reverse hierarchy theory [17,18]. This theory proposes an easy-to-difficult stimulus learning cascade from higher to lower level brain sites, with easy stimulus learning serving as a visual attentional pointer to lower level sites for difficult stimulus learning. Stimulus temporal patterning could serve as such a pointer because it serves to "tag" the stimuli and thus improve top-down selection. Our studies elucidate the type of tagging that the reverse hierarchy theory can make use of for the top-down training of early brain regions. To deepen our understanding of the properties of tagging, in the second part of the study, we conducted a series of experiments aimed at understanding the possible mechanism underlying stimulus temporal patterning. We instantiated the findings in a stimulus tagging model for multi-stimulus perceptual learning.

Results

With stimulus temporal patterning, multiple stimuli are presented in a fixed order with constant inter-trial intervals (ITIs) [3], so that both the stimulus sequence information and a stimulus rhythm are present. To investigate whether learning depends on stimulus sequence alone or on stimulus sequence with a rhythm, we had human observers practice multi-contrast discrimination in rhythmic and non-rhythmic sequence conditions at two ITIs. Four reference contrasts were interleaved in an ascending order (i.e., 0.2, 0.3, 0.47, and 0.63), with ITIs at 2 or 3 s (plus an observer's response time, which was more or less constant across trials), and the ITIs were either constant (2 or 3 s) or jittered (1–3 s or 2–4 s, with the mean at 2 or 3 s). The jittered ITI conditions interrupted the stimulus rhythm but preserved the stimulus sequence information.

Each experimental condition had six observers completing five 2-h practice sessions on different days, and the learning effects (Figure 1C–1F) were compared to constant 1-s fixed

ITI data collected earlier [3]. Significant learning was evident in the 1-s fixed ITI condition (Figure 1B, the post/pre-training threshold ratio (PPR) = 0.69 ± 0.06 ; $PPR < 1$ indicates reduced thresholds after practice and perceptual learning). The results showed that practice with a constant 2-s ITI produced about the same amount of significant learning (PPR = 0.71 ± 0.03 , Figure 1C; F and p values are presented in the figure legends) as did practice with a 1-s ITI, but increasing the constant ITI to 3 s significantly reduced learning ($p = 0.009$), although learning was still significant (PPR = 0.86 ± 0.02 , Figure 1E). Jittering the ITI (for both 2 and 3 s) disabled learning (PPR = 0.99 ± 0.05 for both conditions, Figure 1D and 1F). These results suggest that for learning to occur, stimuli must be delivered in a rhythmic sequence, preferably with closer temporal proximity.

Moreover, we found that evenly spaced stimulus rhythms are most efficient for enabling perceptual learning. For the same stimulus sequence, if the rhythm was made up of uneven ITIs (i.e., 2-, 2.5-, 1.5-, and 1-s ITIs after reference contrasts of 0.2, 0.3, 0.47, and 0.63, respectively), our observers did not learn at all after 5 d of practice (PPR = 1.05 ± 0.05 , Figure 1G). However, if the rhythm was more predictable (i.e., ITI lengthening from 1.25, 1.75, 2.25, to 2.75 s after reference contrasts of 0.2, 0.3, 0.47, and 0.63, respectively), learning was evident again (PPR = 0.82 ± 0.06 , Figure 1H), although marginally weaker than that in the evenly spaced rhythm condition (Figure 1B; $p = 0.105$).

We also compared within- and between-session learning data under constant and jittered 2-s ITI conditions to study the dynamics of perceptual learning. Within-session learning was defined by $Th(5n)/Th(5n-4)$, the ratio of 5th-run threshold over 1st-run threshold in the same session, where n was the session number, and between-session learning was defined by $Th(5n+1)/Th(5n)$, the ratio of 1st-run threshold in the next session over 5th-run threshold in the current session. With rhythmic stimulus sequence (Figure 1J), the mean within-session learning index was 0.85, or on average a 15% threshold decrease within each training session. However, the between-session learning index was 1.04, suggesting no further improvement during the inter-session periods (typically 1–3 calendar days). When the stimulus rhythm was interrupted by jittered ITI (Figure 1K), the within-session learning index was 1.03 (the between-session index was 0.97), suggesting interruption of within-session learning.

The effect of trial-by-trial stimulus roving during practice suggests that the encoding of stimulus information is interrupted. However, it is unclear whether stimulus roving interferes with consolidation of perceptual learning, during which period the stimulus traces transform from working memory to long-term memory. There is evidence that consolidation of perceptual learning is mainly accomplished in the intervals between two consecutive training sessions [19–21]. To examine whether consolidation is subject to roving interference, and if so for how long, we asked whether roving would disrupt learning after temporal-patterned practice.

Twenty-four observers participated in this 5-d experiment. On each day, the observers first practiced five runs of contrast discrimination, each run containing four interleaved stair-



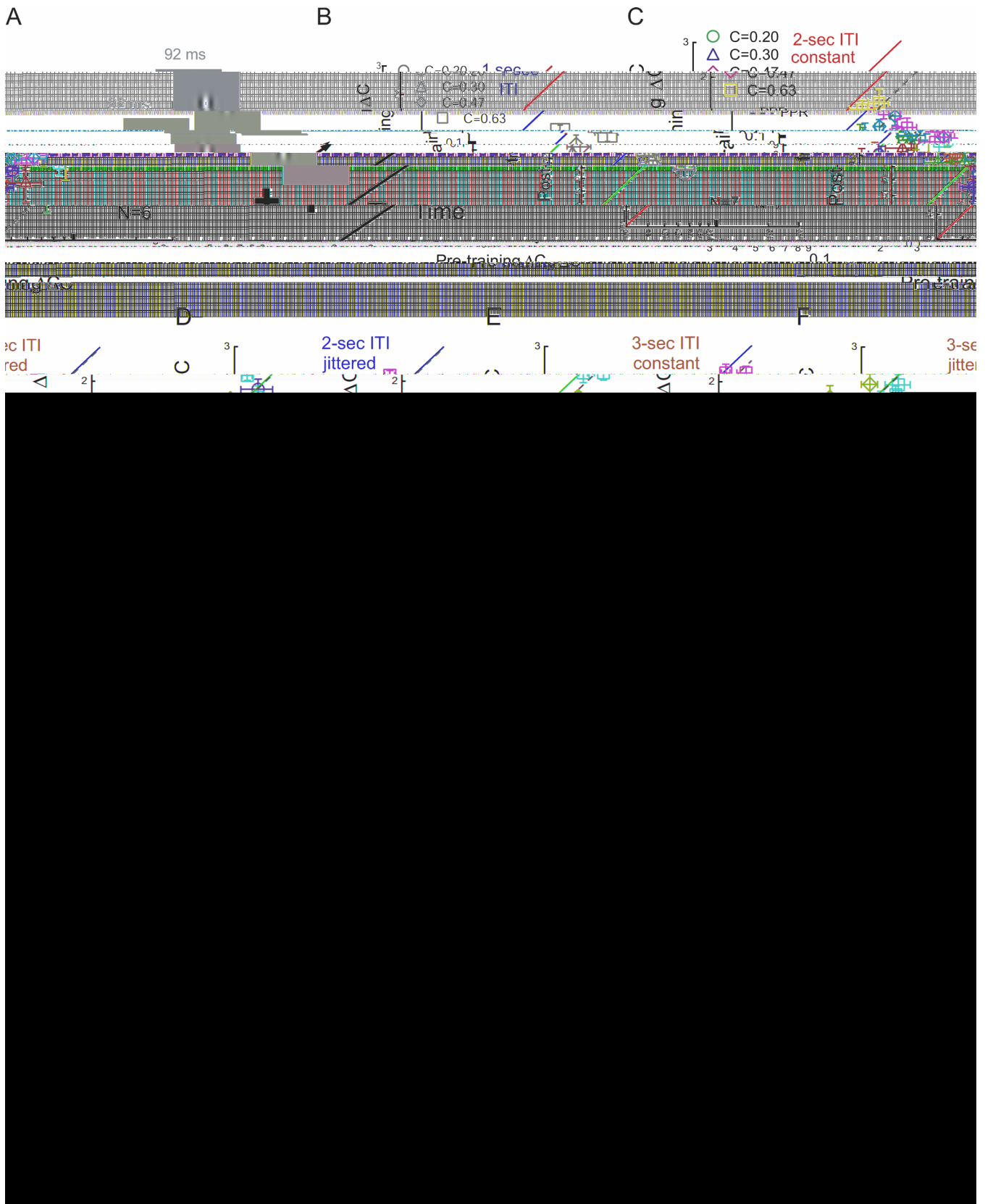


Figure 1.

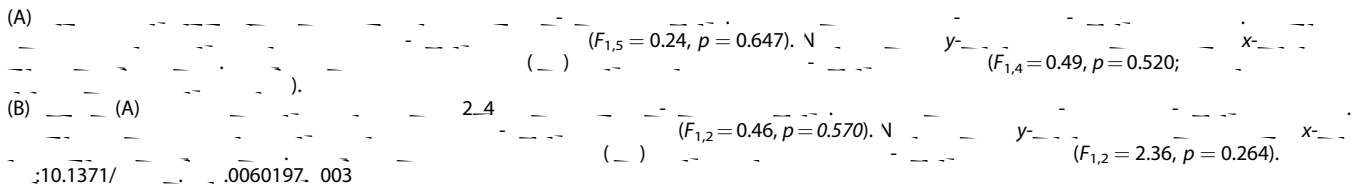
(A) - 2A
 (B) - 1- 3-
 (C) - 2- (F_{1,5} = 89.9, p < 0.001; A) (F_{1,5} = 0.08, p = 0.789)



A



Figure 3.



at least 4 h after each practice session. After that, learning is largely consolidated and fairly immune to roving interruption.

In the following experiments, we investigated whether roving also interrupts stimulus retrieval after successful completion of temporal-patterned training.

In the first experiment, 1 d after five sessions of temporal-patterned training, which produced significant learning (PPR = 0.71 ± 0.03 , left panel of Figure 3A, replotted from Figure 1C for the same observers), six trained observers performed four roving sessions of contrast discrimination for the same reference contrasts. Contrast thresholds in the first roving session did not differ significantly from the post-training thresholds, with a mean 1st-day roving/post-training threshold ratio of 1.02 ± 0.03 (middle panel of Figure 3A). This result indicates a complete transfer of learning from the

trained temporal-patterned condition to the roving condition immediately after completion of training, consistent with our earlier data [3]. Moreover, four sessions of roving runs had no significant impact on learned performance either in five observers (4th-day roving/post-training threshold ratio = 1.05 ± 0.07 , right panel of Figure 3A the sixth observer did not complete all roving sessions). For these five observers the 4th-day roving/pre-training threshold ratio was 0.78 ± 0.07 , similar to their post-/pre-training threshold ratio (0.74 ± 0.04), indicating that learning was not perturbed by extended roving interference.

The second experiment was identical to the first, except for a 2–4-wk gap between the original temporal-patterned training and the four roving sessions. Three new observers completed this experiment. Their initial temporal-patterned training resulted in a PPR of 0.53 ± 0.06 (left panel of Figure 3B). Again there were no significant differences between post-training thresholds and first roving session thresholds (1st-



day roving/post-training threshold ratio = 1.12 ± 0.14 , middle panel of Figure 3B), or between post-training thresholds and last (fourth) roving session thresholds (post-roving/post-training threshold ratio = 1.10 ± 0.17 , right panel of Figure 3B). Learning was unperturbed by four roving sessions, with a 4th-day roving/pre-training ratio of 0.58 ± 0.08 . Results from these two experiments suggest that, once learned, the stimulus traces remain stable and immune to interference by multi-session roving after training; rather, they can be used to guide contrast discrimination in the roving condition.

“block,” so that stimulus traces can build up to establish the stimulus identity and resist roving disruption.

At one extreme, when multiple contrasts are roved from trial to trial, no perceptual learning occurs (Figure 4D, one-trial block size data) [6]. At the other extreme, practice in blocked trials (about 30 trials per block or staircase) produces significant learning (Figure 4D, 30-trial block size data) [6]. To determine the smallest block size for perceptual learning of multiple stimuli, we again had observers practice discrimination of four roving contrasts for five sessions, but this time each roving contrast was roved every three, five or eight consecutive trials. The mean PPR over four contrasts was 0.95 ± 0.07 for the roving-every-three-trial condition (Figure 4A

“B” ?

It is well documented that effective perceptual learning of multiple stimuli can occur when each stimulus (or stimulus level) is practiced in separate blocks [6,12,23]. Our working hypothesis is that when practicing discrimination with several confusable stimuli, each stimulus needs to be tagged so that the brain can attend to the appropriate perceptual template to enable learning (see next experiment and Discussion). When only a single stimulus is practiced in a block, there is no uncertainty about the stimulus tag. What is unknown is how many consecutive trials are necessary to form an effective

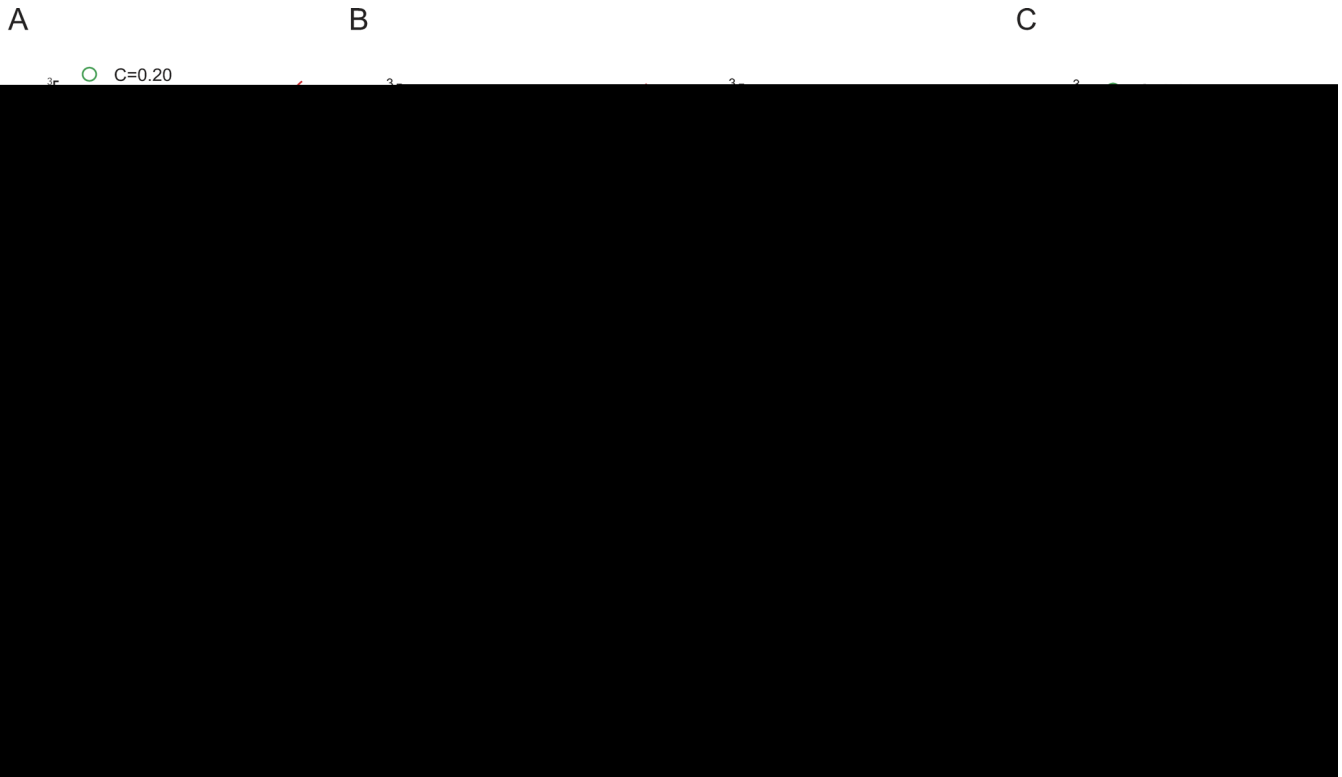


Figure 5. (A) $(F_{1,3} = 3.29, p = 0.167)$. (B) $(0.20, 0.47; F_{1,5} = 23.2, p = 0.005)$. $(0.30, 0.47; F_{1,5} = 0.79, p = 0.414)$. $(F_{1,5} = 34.7, p = 0.002)$. $(F_{1,5} = 27.7, p = 0.003)$. $(F_{1,5} = 40.8, p = 0.001)$. $(F_{1,5} = 0.61, p = 0.472)$. ;10.1371/ .0060197. 005

Why is perceptual learning disabled by roving but enabled by temporal patterning? One way to understand this is to seek exceptions in which perceptual learning succeeds with stimulus roving. Our initial attempts failed to find such exceptions. For example, lengthening the stimulus interval from 92 ms to 400 ms, which reduced stimulus uncertainty, produced no significant performance change ($PPR = 0.91 \pm 0.06$, Figure 5A). Using a spatial rather than a temporal two-alternative forced choice (2AFC) paradigm, which excluded the requirement of working memory for discrimination ($PPR = 1.02 \pm 0.08$, unpublished data), and providing a physical pre-cue identical to the reference stimulus ($PPR = 0.88 \pm 0.10$) [3], produced no significant change either. However, we did find two cases in which roving did not disable multiple-level contrast learning.

First, we had two groups of observers practice discrimination of two roving contrasts, with one group practicing the more similar pair, 0.30 and 0.47, and the other group practicing the less similar pair, 0.20 and 0.47. Perceptual learning of 0.47 contrast does not transfer to 0.30 contrast [6], so these two contrasts, though close, must be processed by independent mechanisms. After five sessions of practice,

observers who practiced the more similar 0.30 and 0.47 pair showed no evidence for significant learning ($PPR = 0.94 \pm 0.07$, left panel of Figure 5B), whereas those who practiced the less similar 0.2 and 0.47 pair showed significant learning ($PPR = 0.76 \pm 0.05$, right panel of Figure 5B). It is interesting that roving just two close stimulus conditions was sufficient to disable perceptual learning. Similar roving effects on two stimuli have also been reported in a bisection learning task [13]. These results indicate that when the two stimuli are very different from each other (probably more than the difference of neighboring mechanisms), roving does not disrupt learning.

Second, the same four roving contrasts (0.2, 0.3, 0.47, and 0.63) were each assigned a tag—the letter A, B, C, and D, respectively. The letters provide ordinal information that could serve to identify the four roving contrasts. In each trial, the letter corresponding to one of the roving contrasts was presented on the screen 200 ms before the onset of first stimulus interval, for 200 ms. Surprisingly, letter cueing restored significant learning during roving after five sessions of training ($PPR = 0.77 \pm 0.04$, Figure 5C). This result is all the more surprising because the letter cue is semantic, whereas previously we found that providing a direct sensory pre-cue (a Gabor patch identical to the reference Gabor) does



not restore much learning [3] It appears that observers need to know the identities of the stimuli to successfully learn the roving stimuli.

To test the generality of these results, we conducted learning experiments with an illusory line orientation discrimination task (far left panel of Figure 5D) to search for conditions in which observers could learn with stimulus roving. Two groups of observers first practiced discrimination of four illusory line orientations (36° , 72° , 108° , and 144°), either with roving or with clockwise rotating temporal patterning. Significant learning was evident with orientation temporal patterning ($PPR = 0.68 \pm 0.06$, middle left panel of Figure 5D), but was absent with orientation roving ($PPR = 0.95 \pm 0.06$, middle right panel of Figure 5D). These data demonstrate the generality of the roles of stimulus roving and temporal patterning in perceptual learning of multiple stimuli. Moreover, unlike stimulus contrast, for which it is difficult to “know” the absolute value, some orientations, i.e. the cardinal and oblique orientations, can be judged with high confidence, and are “known” to observers and not easily confused with other orientations. Therefore, if learning multiple stimuli depends on how well the observers know the stimulus identities, as suggested in the letter cueing experiment, practice would be expected to improve discrimination of these distinct orientations even with roving. To test this notion, we had six new observers practice the four cardinal and oblique orientations with roving, and indeed we found significant learning after five sessions of training ($PPR = 0.66 \pm 0.05$, far right panel of Figure 5D).

Taken together, these results provide important hints about the potential role of stimulus temporal patterning in multiple-stimulus learning: its sequence and rhythm assigns identities to stimuli that would otherwise be confused during learning.

Discussion

Our results demonstrate that a rhythmic stimulus sequence is required to enable perceptual learning. Rhythmic stimulus presentation, especially with evenly spaced trials, would allow the observer to accurately switch attention to the outputs of the most appropriate set of neurons when the observer learns multi-level stimuli (see Figure 5 and our stimulus tagging model below). Jones et al. [24,25] proposed that attention in auditory perception is an inherently oscillatory process with adjustable periodic pulses, and that a system’s responses reach maximal accuracy when attention pulses synchronize with the rhythm of the external stimuli. Our results might be consistent with an oscillation case for attention in visual perception when multi-level stimuli are being learned. Perceptual learning was significantly reduced when the ITI length changed from 2 to 3 s for the current multi-contrast discrimination task, which implies a limited length of the optimal oscillation period.

We also found that the first few hours of learning consolidation after each practice session are subject to disruption by roving. Seitz et al. [20] recently reported that consolidation after a block of training of a stimulus was interrupted if followed immediately by another block of training of a similar but different stimulus. But consolidation was little affected if the second block of training was

conducted 1 h later. Although we and Seitz et al. [20] both study the effects of stimulus roving on perceptual learning, we did not find that consolidation was disrupted by roving.

also to categorize the pitch appropriately as, for example, a knuckleball, curve ball, slider, or fast ball.

Stimulus temporal patterning may similarly aid learning by tagging each roving stimulus according to its “what” (sequence) and “when”

4. Lu ZL, Doshier BA (2004) Perceptual learning retunes the perceptual template in foveal orientation identification. *J Vis* 4: 44–56.
5. Saarinen J, Levi DM (1995) Perceptual learning in vernier acuity: What is learned? *Vision Res* 35: 519–527.
6. Yu C, Klein SA, Levi DM (2004) Perceptual learning in contrast discrimination and the (minimal) role of context. *J Vis* 4: 169–182.
7. Fine I, Jacobs RA (2002) Comparing perceptual learning tasks: a review. *J Vis* 2: 190–203.
8. Fahle M (2005) Perceptual learning: specificity versus generalization. *Curr Opin Neurobiol* 15: 154–160.
9. Stickgold R, Walker MP (2005) Memory consolidation and reconsolidation: what is the role of sleep? *Trends Neurosci* 28: 408–415.
10. Walker MP, Brakefield T, Hobson JA, Stickgold R (2003) Dissociable stages of human memory consolidation and reconsolidation. *Nature* 425: 616–620.
11. Nader K (2003) Memory traces unbound. *Trends Neurosci* 26: 65–72.
12. Adini Y, Wilkowsky A, Haspel R, Tsodyks M, Sagi D (2004) Perceptual learning in contrast discrimination: the effect of contrast uncertainty. *J Vis* 4: 993–1005.
13. Otto TU, Herzog MH, Fahle M, Zhaoping L (2006) Perceptual learning with spatial uncertainties. *Vision Res* 46: 3223–3233.
14. Adini Y, Sagi D, Tsodyks M (2002) Context-enabled learning in the human visual system. *Nature* 415: 790–793.
15. Tsodyks M, Adini Y, Sagi D (2004) Associative learning in early vision. *Neural Netw* 17: 823–832.
16. Doshier BA, Lu ZL (1998) Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc Natl Acad Sci U S A* 95: 13988–13993.
17. Ahissar M, Hochstein S (1997) Task difficulty and the specificity of perceptual learning. *Nature* 387: 401–406.
18. Ahissar M, Hochstein S (2004) The reverse hierarchy theory of visual perceptual learning. *Trends Cogn Sci* 8: 457–464.
19. Karni A, Tanne D, Rubenstein BS, Askenasy JJ, Sagi D (1994) Dependence on REM sleep of overnight improvement of a perceptual skill. *Science* 265: 679–682.
20. Seitz AR, Yamagishi N, Werner B, Goda N, Kawato M, et al. (2005) Task-specific disruption of perceptual learning. *Proc Natl Acad Sci U S A* 102: 14895–14900.
21. Mednick S, Nakayama K, Stickgold R (2003) Sleep-dependent learning: a nap is as good as a night. *Nat Neurosci* 6: 697–698.
22. Censor N, Karni A, Sagi D (2006) A link between perceptual learning, adaptation and sleep. *Vision Res* 46: 4071–4074.
23. Ahissar M, Laiwand R, Kozminsky G, Hochstein S (1998) Learning pop-out detection: building representations for conflicting target-distractor relationships. *Vision Res* 38: 3095–3107.
24. Large EW, Jones MR (1999) The dynamics of attending: how people track time-varying events. *Psychol Rev* 106: 119–159.
25. Jones MR, Moynihan H, MacKenzie N, Puente J (2002) Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychol Sci* 13: 313–319.
26. McGaugh JL (2000) Memory—a century of consolidation. *Science* 287: 248–251.
27. Parkosadze K, Otto TU, Malania M, Kezeli A, Herzog MH (2007) Perceptual learning of bisection stimuli under roving: slow and largely specific. *J Vis* 8: 1–8.

