Identifying Long- and Short-Term Processes in Perceptual Learning

Jia Yang[1,2], Fang-Fang Yan[1,2], Lijun Chen[1,2], Shuhan Fan[1,2], Yifan Wu[1,2], Lei Jiang[1,2], Jie Xi[1,2], Junlei Zhao[3,4], Yudong Zhang[3,4], Zhong-Lin Lu[5,6,7,8], and Chang-Bing Huang[1,2]

[1]Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences; [2]Department of Psychology, Chinese Academy of Sciences; [3]Key Laboratory of Adaptive Optics, Chinese Academy of Sciences; [4]Institute of Optics and Electronics, Chinese Academy of Sciences; [5]Division of Arts and Sciences, New York University Shanghai; [6]Center for Neural Science, New York University; [7]Department of Psychology, New York University; and [8]NYU-ECNU Institute of Brain and Cognitive Science at New York University Shanghai

Abstract

Practice makes perfect in almost all perceptual tasks, but how perceptual improvements accumulate remains unknown. Here, we developed a multicomponent theoretical framework to model contributions of both long- and short-term processes in perceptual learning. Applications of the framework to the block-by-block learning curves of 49 adult participants in seven perceptual tasks identified ubiquitous long-term general learning and within-session relearning in most tasks. More importantly, we also found between-session forgetting in the vernier-offset discrimination, face-view discrimination, and auditory-frequency discrimination tasks; between-session off-line gain in the visual shape search task; and within-session adaptation and both between-session forgetting and off-line gain in the contrast detection task. The main results of the vernier-offset discrimination and visual shape search tasks were replicated in a new experiment. The multicomponent model provides a theoretical framework to identify component processes in perceptual learning and a potential tool to optimize learning in normal and clinical populations.

Perceptual learning often improves performance through thousands of trials of practice across multiple days (Karni & Sagi, 1993). Although long-term benefits of perceptual learning (Sagi, 2011; Watanabe & Sasaki, 2015) have been observed in a wide range of tasks, from simple feature detection (Fine & Jacobs, 2002) to complex object discrimination (Frank et al., 2020), a number of short-term phenomena, including within-session adaptation (Censor et al., 2006), between-session consolidation (Karni et al., 1994; Stickgold et al., 2000; Tamaki et al., 2020; Yotsumoto et al., 2009), and off-line gain (Bang et al., 2018; Shibata et al., 2017), have also been documented in some tasks. However, other well-known phenomena such as between-session forgetting (De Groot & Keijzer, 2000; Pavlik & Anderson, 2005; Storm et al., 2011) and faster relearning (Jaber, 2005), well documented in other learning domains such as language learning (De Groot & Keijzer, 2000) and skill acquisition (Rickard, 2007), have not been observed in perceptual learning (but see Beard et al., 1995; Mascetti et al., 2013). The contributions of long- and short-term processes in perceptual learning remain a fundamental question in studies of perceptual learning.

Corresponding Authors:
Zhong-Lin Lu, New York University Shanghai, Division of Arts and Sciences
Email: zhonglin@nyu.edu

Chang-Bing Huang, Chinese Academy of Sciences, Institute of Psychology, Key Laboratory of Behavioral Science
Email: huangcb@psych.ac.cn

In this study, we developed a multicomponent theoretical framework (Fig. 1) to model and identify contributions of both long- and short-term processes to perceptual learning and applied it to an existing data set with 49 adult participants trained in seven different tasks and another experiment with 14 adult participants trained in two tasks. We hypothesized that improved perceptual-task performance through training reflects cumulative effects of both long- and short-term processes that can be revealed by analyzing the fine-grained block-by-block (e.g., tens of trials) learning curves. The common practice in perceptual-learning studies is to sample the learning curve at the session level (Dosher & Lu, 2007; Kattner et al., 2017). The coarse sampling scheme may have obscured some local structures in the learning curve and concealed certain short-term processes (Kattner et al., 2017). In addition, different training tasks may engage different long- and short-term processes. For example, sleep-enabled between-session improvement has been observed in texture discrimination and orientation identification tasks (Mascetti et al., 2013; Stickgold et al., 2000; Tamaki et al., 2020) but not in the chevron- and face-identification tasks (Aberg et al., 2009; Hussain et al., 2008).

The theoretical framework incorporates a number of possible long- and short-term processes, including between-session general learning, off-line gain and forgetting, and within-session relearning and adaptation, and their signature effects in both fine- and coarse-grained analysis. If learning were just a function of training intensity (e.g., number of trials), a single power or exponential function (Fig. 1f), representing cumulative effects of neural plasticity, would suffice to describe the learning curve regardless of how training was distributed across days (Hussain et al., 2009). In contrast, phased learning curves with increasing or decreasing learning rates across sessions (Figs. 1g–1j) would indicate additional processes (Levi et al., 1997). In this framework, general learning and within-session relearning represent benefits from repeated practice, whereas within-session adaptation is used to model detrimental effects of repeated exposure (Censor et al., 2006; Levi et al., 1997). We also introduced between-session components to model performance loss (forgetting; Figs. 1b and 1g; Beard et al., 1995; Levi et al., 1997) or improvement (off-line gain; Figs. 1c and 1h; Karni et al., 1994; Tamaki et al., 2020) in the beginning of each new session.

We first applied the theoretical framework to analyze the fine-grained block-by-block learning curves in an existing data set consisting of 49 participants trained in seven perceptual tasks across 35 sessions. This data set was collected to examine individual differences,

Statement of Relevance

How repeated training or practice leads to long-term performance improvements is one of the fundamental questions in skill acquisition. Typically, performance improves during or immediately following each training session, but the performance improvements may either increase through consolidation or diminish because of forgetting between training sessions. On the other hand, training over multiple sessions often leads to overall performance improvements that may last a long time after training. In this study, we developed a theoretical framework to identify the various component processes in perceptual learning and explain the detailed time course of performance change across multiple training sessions. We applied the new framework to model the block-by-block learning curves of 49 adult participants in seven perceptual-learning tasks. We found ubiquitous long-term general learning and within-session relearning in most tasks and between-session gain, between-session forgetting, and within-session adaptation in some tasks. The theoretical framework provides not only a tool to identify component processes in perceptual learning but also a potential tool to optimize learning outcomes in normal and clinical populations.

functional forms of the learning curves, and effects of training sequence in perceptual learning. An article on individual differences was published in 2020 (Yang et al., 2020). In this study, applications of the proposed framework successfully identified long-term general learning and within-session relearning in most tasks and between-session forgetting and off-line gain, as well as within-session adaptation, in a subset of the tasks. To exclude the potential influence of the training paradigm and evaluate generalizability of the results, we conducted a new experiment with 14 new participants who were trained only with vernier-offset discrimination and visual shape search tasks. The main results were replicated in these two tasks.

Method

Participants

Forty-nine participants (22 male; age: M = 23.4 years, range = 10 years) were recruited from the Institute of Psychology at the Chinese Academy of Sciences or from
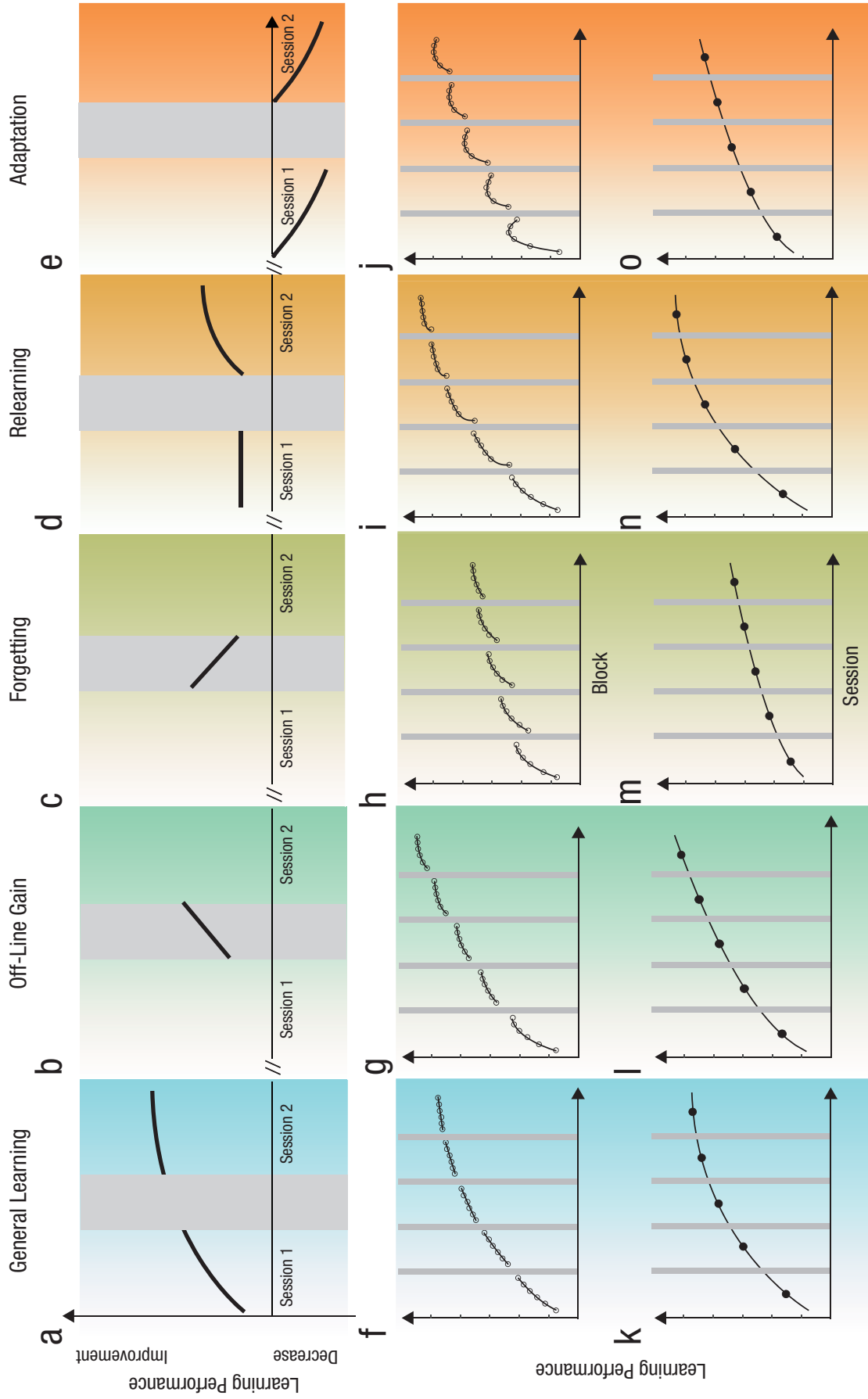
Fig. 1. Long- and short-term processes and their signature effects on learning curves across sessions. The top row shows effects of (a) general learning, (b) off-line gain, (c) off-line forgetting, (d) relearning, and (e) adaptation within or between two consecutive sessions. The middle row shows signature effects of the long- and short-term processes (f–j) on multisession learning curves in a fine-grained block-by-block analysis. The bottom row shows effects of the long- and short-term processes (k–o) on multisession learning curves in a coarse-grained session-by-session analysis. Gray bars indicate delays between sessions, open circles indicate block-by-block performance scores, and filled circles indicate session-by-session performance scores.

nearby universities and communities. All participants were undergraduate or graduate students. All were naive to the purpose of the study, were right handed, had normal or corrected-to-normal vision, and had no history of psychiatric or neurological disorders. The Institutional Ethical Committee of the Institute of Psychology approved the study. Prior to data collection, all participants provided written informed consent. A power analysis with G*Power (Version 3.1; Faul et al., 2007) found that a sample size of 49 would be needed to achieve an effect size (f) of 0.2 given an α of .05 and power ($1 - β$) of .96.

Apparatus

The visual and audiovisual n-back working memory tasks were conducted on two Sony G220 monitors. The resolution was 1,600 × 1,200 pixels, and the refresh rate was 85 Hz. The background luminance was 36 cd/m². A 14-bit gray resolution was achieved via a special circuit that combined two 8-bit output channels of the graphics card (Li et al., 2003). For the auditory task, we used a Dell E1912Hc LCD monitor. All stimuli were programmed in MATLAB (The MathWorks, Natick, MA) with Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

Design and procedure

Participants practiced seven tasks—contrast detection, vernier-offset discrimination, motion-direction discrimination, visual shape search, face-view discrimination, auditory-frequency discrimination, and audiovisual n-back working memory tasks (for brief descriptions of the tasks, see Fig. 2; for further details, see the Tasks section in the Supplemental Material available online). The order of training was counterbalanced across participants with a Latin-square design. Participants were trained on each task in five daily sessions that lasted 40 min to 1 hr. The mean temporal interval between two adjacent sessions was 23.06 hr ($SD = 9.37$ hr, $Mdn = 23.09$ hr).

For the vernier-offset discrimination, motion-direction discrimination, face-view discrimination, and auditory-frequency discrimination tasks, participants received 3,500 trials of training, allocated in five sessions of seven blocks of 100 trials each. For the contrast detection and visual shape search tasks, participants received 3,360 trials of training, allocated in five sessions of seven blocks of 96 trials each. For the audiovisual n-back working memory task, participants received 3,000 trials of training, allocated in five sessions of 30 miniblocks that contained 20 + n trials each. In total, each participant received about 23,720 trials (Fig. 2h). Ten to 20 instruction trials were provided before data collection in the first session of each task.

Data analysis

During the experiment, we adopted a three-down/one-up staircase procedure (Levitt, 1971) targeting the 79.4% correct performance level in the contrast detection, vernier-offset discrimination, face-view discrimination, and auditory-frequency discrimination tasks. The adaptive staircase method is widely used to keep the performance level nearly constant in perceptual learning (Amitay et al., 2005; Levitt, 1971; Shibata et al., 2017). To reduce the bias of threshold estimates from the staircase method (Lu & Dosher, 2013), we used a slightly higher stimulus level in the first trial in each session. We calculated the average performance level (percentage of correct responses) across every 48 (contrast detection) or 50 (vernier-offset discrimination, face-view discrimination, and auditory-frequency discrimination) trials (see Fig. S1 in the Supplemental Material). We found that most of the observed performance levels were within ±5% of the targeted 79.4% level. Thresholds in each block and session were then estimated using a maximum likelihood procedure (see the Models and Analysis section in the Supplemental Material; Myung, 2003), and perceptual performance in each training block was defined as the reciprocal of the maximum-likelihood-estimation-based thresholds at 79.4% correct. For the motion-direction discrimination and visual shape search tasks, we used the method of constant stimuli, and the measured percentage of correct responses in each block was converted into d'. For the audiovisual n-back working memory task, the n-back level was adjusted adaptively across miniblocks, and the average n-back level from every five miniblocks was used to index the performance in the block. A one-way repeated measures ANOVA was conducted to examine learning effects at both the coarse (session-wise) and fine (block-wise) scale for each task, respectively. We also conducted paired-samples t tests between the first and last blocks within each session and between the last block in one session and the first block in the subsequent session to test within-session learning and between-session effects. We calculated adjusted p values with the Benjamini-Hochberg method to control the false-discovery rate in multiple comparisons.

To simultaneously reveal general learning, between-session off-line gain and forgetting, and within-session adaptation and relearning in the learning process, we fitted a series of models to the average block-by-block learning curves and performed nested model comparisons (Lu & Dosher, 2013) and Bayesian model-selection analysis by computing the posterior probability p_{BIC} of two competing hypotheses to identify the best-fitting model. Bayesian information criterion (BIC) is a criterion for model selection among a finite set of models based on the Bayesian posterior probability of a candidate
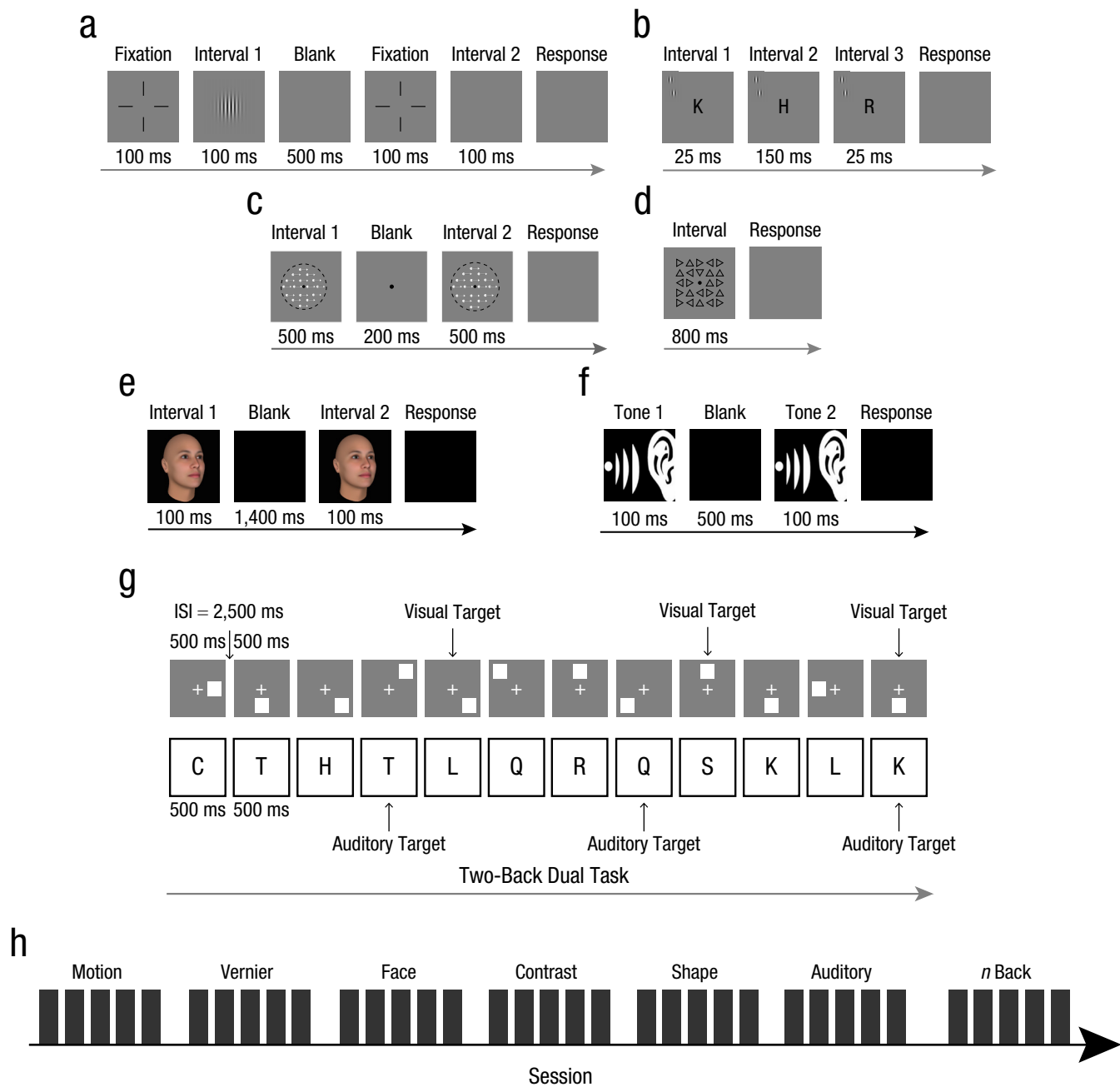
Fig. 2. Illustration of the training tasks and design of the study (figure and study design adapted from Yang et al., 2020). In the contrast detection task (a), participants were required to report which of two intervals a target grating had appeared in. In the vernier-offset discrimination task (b), participants were shown a series of letters in their foveal vision along with a vernier stimulus (two gratings) in their peripheral vision; they were asked to identify both the foveal letter and the offset direction of the vernier stimulus (i.e., whether the lower grating was to the left or right of the upper grating). In the motion-direction discrimination task (c), participants saw a moving stimulus in two separate intervals and had to judge whether the two stimuli were moving in the same direction. In the visual shape search task (d), participants were shown a display containing 24 triangles and had to judge whether one of them (the target) had a downward orientation. In the face-view discrimination task (e), participants were shown a face image in two separate intervals and had to judge whether the second face tilted left or right relative to the first face. In the auditory-frequency discrimination task (f), participants listened to a tone in two separate intervals and had to judge which interval contained the higher tone. In the audiovisual *n*-back working memory task (g), participants were shown a series of visual or auditory images and asked to judge whether the current stimulus matched the one that preceded it by *n* trials in the stimulus stream (e.g., same location or same resonance). Each participant received five sessions of training for each of the seven tasks sequentially; one example training sequence is shown in (h). The training sequence was counterbalanced across participants with a Latin-square design. For more information and detailed descriptions, see the Tasks section in the Supplemental Material available online. ISI = interstimulus interval.

model (Neath & Cavanaugh, 2012). p_{BIC} quantifies the evidence regarding which model (e.g., reduced model vs. full model) is more strongly supported by the data (Masson, 2011; for details, see Models and Analysis in the Supplemental Material). The full model contained 24 parameters, including two parameters for general learning, four for off-line gain and forgetting, four for the magnitude of relearning, four for the rate of relearning within each session, five for the magnitude of adaptation, and five for the rate of adaptation:

$$\hat{T}(j,k) = \alpha\left(5(j-1)+k\right)^{\tau} + s1(j) + s2(j)(k-1)^{\tau_{s2}(j)}$$
$$- s3(j)(k-1)^{\tau_{s2}(j)}$$

$$s1(j) = 0, \ g1, \ g2, \ g3, \ g4 \qquad \text{(for } j = 1 \sim 5)$$

$$s2(j) = 0, \ r1, r2, r3, r4 \qquad \text{(for } j = 1 \sim 5)$$

$$s3(j) = a1, a2, a3, a4, a5 \qquad \text{(for } j = 1 \sim 5)$$

$$\tau_{s2}(j) = 0, \ \gamma1, \ \gamma2, \ \gamma3, \ \gamma4 \qquad \text{(for } j = 1 \sim 5)$$

$$\tau_{s3}(j) = \beta1, \beta2, \beta3, \beta4, \beta5 \qquad \text{(for } j = 1 \sim 5),$$

where j denotes the jth session and k denotes the kth block. $\hat{T}(j,k)$ represents the predicted performance in the kth block of the jth session; α and τ denote the initial performance and learning rate of general learning, respectively; $s1$ denotes the magnitude of off-line gain (positive value) or forgetting (negative value) in four between-session delays; $s2$ and τ_{s2} denote the magnitude of improvement between two adjacent blocks and the rate-of-relearning component in four sessions (from Session 2 to Session 5); and $s3$ and τ_{s3} denote the magnitude of decline between two adjacent blocks and the rate-of-adaptation component in five sessions. Different versions of the reduced model (see Fig. S3 and the Results section in the Supplemental Material) were also fitted to the learning curves.

Results

Learning curves

The average block-by-block learning curves are shown in Figure 3. There was general learning across all the blocks in all tasks, indicating long-term learning (Sagi, 2011; Watanabe & Sasaki, 2015; Figs. 1a, 1f, and 3a–3g). For the visual shape search task, there were evident between-session improvements, consistent with off-line gain (Karni et al., 1994; Stickgold et al., 2000; Tamaki et al., 2020; Figs. 1b, 1g, and 3d). For the vernier-offset discrimination, face-view discrimination, and auditory-frequency discrimination tasks, there was substantial between-session performance deterioration, suggesting

between-session forgetting after each daily training (Figs. 1c, 1h, 3b, 3e, and 3f). For the contrast detection task, there was within-session performance deterioration, consistent with within-session adaptation (Censor et al., 2006; Figs. 1e, 1j, and 3a). For the vernier-offset discrimination, face-view discrimination, and auditory-frequency discrimination tasks, there was rapid within-session learning, indicative of relearning (Figs. 1d, 1i, 3b, 3e, and 3f). None of the rich patterns of results except general learning were observed in the session-by-session analysis (see Fig. S2 and Table S2 in the Supplemental Material).

To confirm these observations, we conducted a series of statistical tests on the learning curves. To evaluate general learning across sessions, we conducted a regression analysis using the power function with two parameters on session-by-session learning curves (see Fig. S2), repeated measures ANOVAs on performance scores across five sessions, and paired-samples t tests between consecutive sessions. We found a significant main effect of session in all tasks ($p < 10^{-10}$, $\eta_p^2 = .27 \sim .83$, corrected for multiple comparisons ($N = 7$) with the Benjamini-Hochberg method; for detailed η_p^2 and 95% CIs for each task, see Table S1 in the Supplemental Material), indicating significant between-session learning. We also conducted a one-way repeated measures ANOVA on performance scores across the seven blocks in each session of each task to examine within-session learning. The results showed that the main effect of block was significant for almost all sessions in six of the seven tasks (25 of 30 comparisons, $p < .05$, $\eta_p^2 = .05 \sim .32$, corrected for 35 comparisons with the Benjamini-Hochberg method; for details, see Table S3 in the Supplemental Material). The contrast detection task was the exception—participants exhibited no within-session learning ($p > .36$ for all five sessions, $\eta_p^2 = .003 \sim .02$; see Table S3), indicative of within-session adaptation (Fig. 1e). The results were also supported by paired-samples t tests between performance scores in the first and last blocks of the same session (Fig. 4; see Table S4 in the Supplemental Material). After correction for multiple comparisons ($N = 63$), the effect of within-session learning (the first block vs. the last block in the same session) was significant in most sessions of all tasks except contrast detection.

To reveal between-session effects, we conducted paired-samples t tests between performance scores in the last block of each session and the first block of the subsequent session (Fig. 4; also see Table S5 in the Supplemental Material). We found significant performance deterioration in the first block of three sessions of the auditory-frequency discrimination task ($p = .04$, Cohen's $d = 0.40$; $p < .001$, Cohen's $d = 0.68$; $p < .02$, Cohen's $d = 0.42$ in the second, third, and fourth sessions, respectively, after correction for multiple
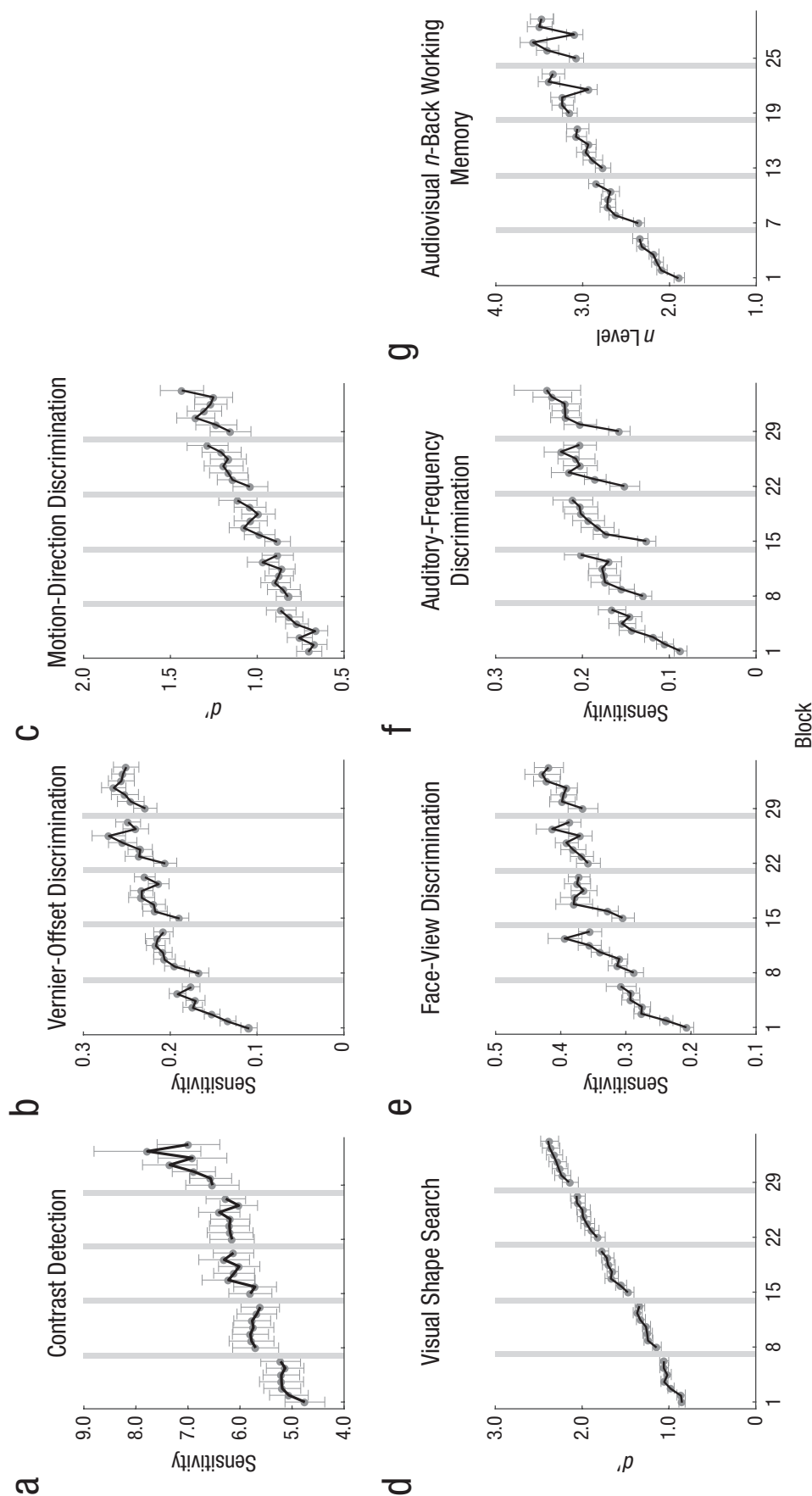
Fig. 3. Block-by-block learning curves in the (a) contrast detection, (b) vernier-offset discrimination, (c) motion-direction discrimination, (d) visual shape search, (e) face-view discrimination, (f) auditory-frequency discrimination, and (g) audiovisual *n*-back working memory tasks. Gray bars indicate between-session delays. Error bars represent standard errors.
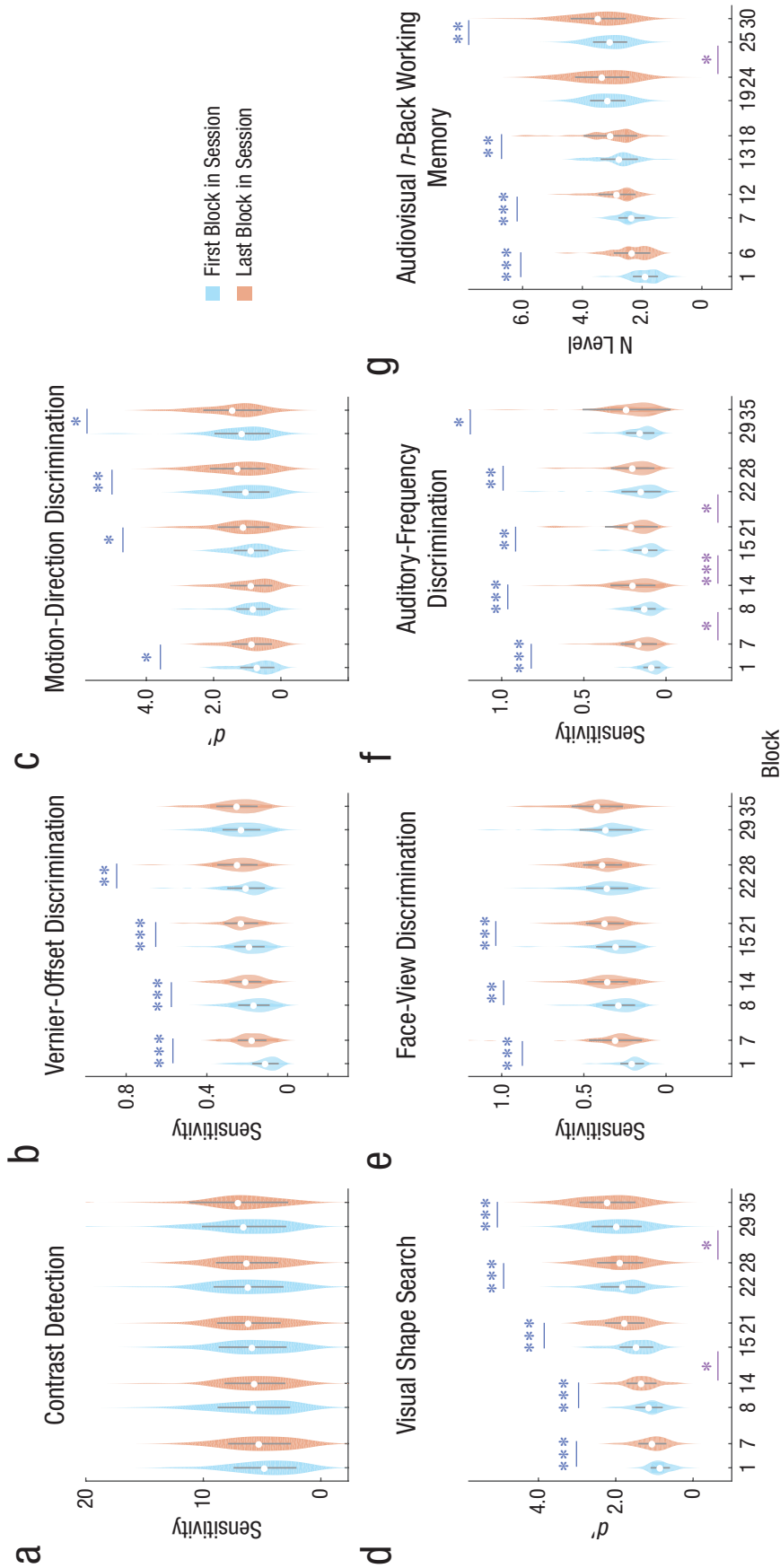
Fig. 4. Within- and between-session effects in the (a) contrast detection, (b) vernier-offset discrimination, (c) motion-direction discrimination, (d) visual shape search, (e) face-view discrimination, (f) auditory-frequency discrimination, and (g) audiovisual *n*-back working memory tasks. The distributions of all 49 participants' performance scores in the first (light blue) and last (light orange) block of each session are visualized by violin plots; white dots represent mean performance in the corresponding block, and error bars represent standard deviations. Blue and purple asterisks represent significant within- and between-session effects, respectively (*$p < .05$, **$p < .01$, ***$p < .001$).

comparisons) and the fifth session of the audiovisual *n*-back working memory task ($p = .03$, Cohen's $d = 0.35$), indicating between-session forgetting after daily training. We also found a tendency for decline in the beginning of a new session for the face-view discrimination task ($p = .06$, Cohen's $d = 0.41$ in the third session after correction for multiple comparisons). In contrast, significant between-session improvements were found in the first block of the third and fifth sessions of the visual shape search task ($p = .03$, Cohen's $d = 0.32$ and $p = .03$, Cohen's $d = 0.14$, respectively, after correction for multiple comparisons). To exclude the influence of the length of the between-session delay on between-session effects, we conducted a two-way mixed ANOVA with task as a between-subject factor (seven tasks) and delay as a within-subject factor (four between-session delays). We found no significant main effects of task, $F(6, 288) = 1.39$, $p = .22$, $\eta_p = .02$, and delay, $F(3, 144) = 2.03$, $p = .11$, $\eta_p = .01$, and no significant interaction between them, $F(18, 864) = 1.27$, $p = .20$, $\eta_p = .02$.

Identifying long- and short-term processes in perceptual learning

To identify the long- and short-term processes in perceptual learning, we fitted a series of models (for details, see the Data Analysis section), computed the goodness of fit (r^2) with a nonlinear least-squares method (see Equation 6 in the Supplemental Material) for each model, and performed model comparisons (*F* test; Lu & Dosher, 2013) and a Bayesian model-selection procedure (Masson, 2011; see the Models and Analysis section in the Supplemental Material) to identify the best-fitting model.

The most saturated model had five components (Figs. 1a–1e): (a) general cross-block learning, (b) between-session off-line gain, (c) between-session off-line forgetting, (d) within-session relearning, and (e) within-session adaptation. Models made of a subset of the five components were also fitted to the data. The best-fitting model for each task was defined as the one that was statistically equivalent to the most saturated model, was superior to all its reduced versions, and had the least number of parameters. In addition, we estimated the standard errors of the parameters of the best-fitting models with a bootstrap procedure (see the Models and Analysis section in the Supplemental Material; Lu & Dosher, 2013).

The best-fitting models for all tasks included general learning across all the blocks (Fig. 5; also see Fig. S3 and Table S6 in the Supplemental Material). For the contrast detection task, the best-fitting model also included within-session adaptation and relearning as well as between-session gain (Fig. 5a), $r^2 = .934$; versus the most reduced model: $r^2 = .772$, $F(9, 24) = 6.55$, $p < .001$, $p_{\text{BIC}}\left(M_{\text{reduced}}|D\right) = .006$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) = .994$; versus

the full model: $r^2 = .949$, $F(13, 11) = 0.25$, $p > .99$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) > .999$, $p_{\text{BIC}}\left(M_{\text{full}}|D\right) < .001$. The best-fitting models included constant within-session relearning and between-session forgetting (Figs. 5b, 5e, and 5f) for vernier-offset discrimination—$r^2 = .951$; versus the most reduced model: $r^2 = .902$, $F(3, 30) = 10.23$, $p < .001$, $p_{\text{BIC}}\left(M_{\text{reduced}}|D\right) = .002$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) = .998$; versus the full model: $r^2 = .972$, $F(19, 11) = 0.44$, $p = .95$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) > .999$, $p_{\text{BIC}}\left(M_{\text{full}}|D\right) < .001$; face-view discrimination—$r^2 = .938$; versus the most reduced model: $r^2 = .886$, $F(3, 30) = 8.49$, $p < .001$, $p_{\text{BIC}}\left(M_{\text{reduced}}|D\right) = .006$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) = .994$; versus the full model: $r^2 = .960$, $F(19, 11) = 0.31$, $p = .99$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) > .999$, $p_{\text{BIC}}\left(M_{\text{full}}|D\right) < .001$; and auditory-frequency discrimination—$r^2 = .952$; versus the most reduced model: $r^2 = .750$, $F(3, 30) = 41.65$, $p < .001$, $p_{\text{BIC}}\left(M_{\text{reduced}}|D\right) < .001$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) > .999$; versus the full model: $r^2 = .975$, $F(19, 11) = 0.52$, $p = .90$, $p_{\text{BIC}}(M_{\text{best}}|D) > .999$, $p_{\text{BIC}}(M_{\text{full}}|D) < .001$. The best-fitting model for the visual shape search task included constant within-session relearning and between-session gain (Fig. 5d)—$r^2 = .997$; versus the most reduced model: $r^2 = .945$, $F(3, 30) = 168.58$, $p < .001$, $p_{\text{BIC}}\left(M_{\text{reduced}}|D\right) < .001$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) > .999$; versus the full model: $r^2 = .998$, $F(19, 11) = 0.56$, $p = .87$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) > .999$, $p_{\text{BIC}}\left(M_{\text{full}}|D\right) < .001$. There was within-session relearning (Figs. 5c and 5g) for motion-direction discrimination—$r^2 = .950$; versus the most reduced model, $r^2 = .872$, $F(2, 31) = 23.85$, $p < .001$, $p_{\text{BIC}}\left(M_{\text{reduced}}|D\right) < .001$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) > .999$; versus the full model, $r^2 = .965$, $F(20, 11) = 0.23$, $p > .99$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) > .999$, $p_{\text{BIC}}\left(M_{\text{full}}|D\right) < .001$, and audiovisual *n*-back working memory—$r^2 = .943$; versus the most reduced model, $r^2 = .927$, $F(2, 26) = 3.68$, $p = .04$, $p_{\text{BIC}}\left(M_{\text{reduced}}|D\right) = 0.445$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) = 0.555$; versus the full model, $r^2 = .964$, $F(20, 6) = 0.18$, $p > .99$, $p_{\text{BIC}}\left(M_{\text{best}}|D\right) > .999$, $p_{\text{BIC}}\left(M_{\text{full}}|D\right) < .001$.

We also fitted the learning curves of each individual participant. The results were largely consistent with those from the group-average analysis (see Fig. S4 in the Supplemental Material). The magnitude of relearning was negatively correlated with that of between-session effects (gain or forgetting) in the contrast detection, vernier-offset discrimination, face-view discrimination, visual shape search, and auditory-frequency discrimination tasks (see Table S7 in the Supplemental Material). There was no significant correlation between within-session adaptation, which was identified only in the contrast detection task, and relearning or between-session gain or forgetting (see Table S7).

To examine whether the observed long- and short-term effects were specific to the settings of the experiment (i.e., seven consecutive training tasks), we conducted a new experiment and trained 14 adult participants (seven male; age: $M = 20.5$ years, range = 10 years) in the visual shape search and vernier-offset
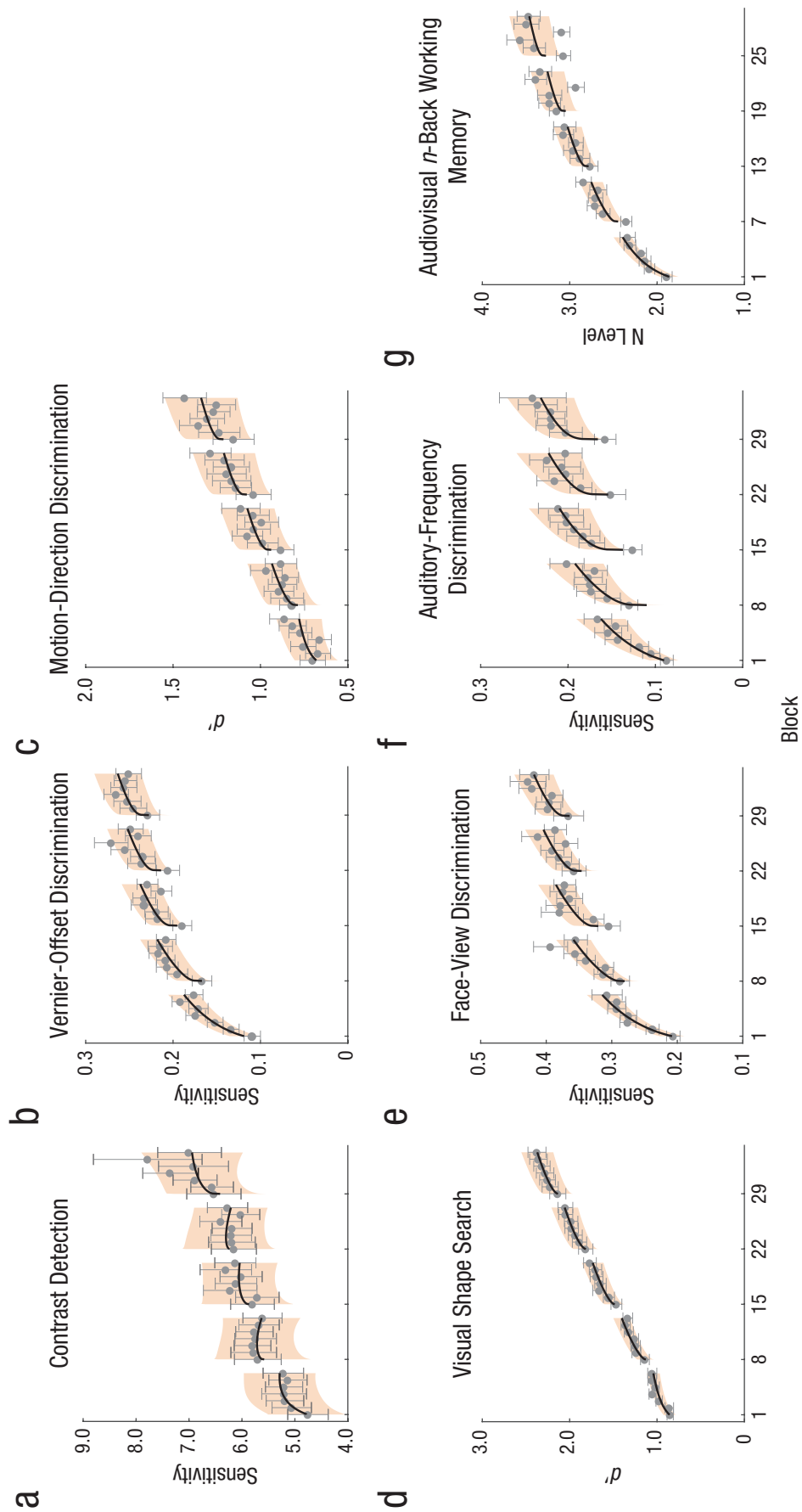
Fig. 5. Predicted learning curves from the best-fitting models in the (a) contrast detection, (b) vernier-offset discrimination, (c) motion-direction discrimination, (d) visual shape search, (e) face-view discrimination, (f) auditory-frequency discrimination, and (g) audiovisual *n*-back working memory tasks. Dots represent measured performance scores. Error bars represent standard errors. Colored bands represent 95% confidence intervals.

discrimination tasks (for details, see the Supplemental Material). We fitted the learning curves in each group in each task as well as the average learning curves from all 14 participants. The results were largely consistent with those of the 49 participants, except that we observed some additional within-session adaptation (see Fig. S6 and the Results section in the Supplemental Material), which might have been caused by the long training sessions (768 trials per session). These results provided compelling evidence for the reliability and generalizability of our findings.

Discussion

In this study, we developed a multicomponent theoretical framework to systematically model and identify both long- and short-term processes in perceptual learning and applied the framework to analyze the fine-grained block-by-block learning curves of 49 participants trained in seven perceptual tasks across 35 sessions. Whereas ubiquitous long-term general learning was observed in all seven tasks, different short-term processes were found: within-session relearning in all tasks; between-session forgetting in the vernier-offset discrimination, face-view discrimination, and auditory-frequency discrimination tasks; between-session off-line gain in the visual shape search task; and within-session adaptation and both between-session forgetting and off-line gain in the contrast detection task. The main results in the vernier-offset discrimination and visual shape search tasks were replicated in a new experiment with only the two tasks, although some additional within-session adaptation was found because of the longer training sessions in the new experiment. Taken together, these results strongly support the multicomponent theoretical framework and its application in identifying component processes in perceptual learning.

We found compelling evidence for between-session forgetting in several perceptual tasks. Although previous studies examining the time course of learning from the perspective of memory formation have revealed off-line gain in motor learning (Walker et al., 2003) and visual perceptual learning (Bang et al., 2018), between-session forgetting (Ebbinghaus, 2013) has not been observed in perceptual learning (but see signs of worse performance in new sessions in the work by Beard et al., 1995, and Mascetti et al., 2013). This is inconsistent with the well-documented posttraining performance decline in vocabulary learning (De Groot & Keijzer, 2000; Pavlik & Anderson, 2005) and skill acquisition (Jaber, 2005; Storm et al., 2011) as well as brain-wide activity decay that was postulated to systematically remove selected memories (Draguhn, 2018; Hardt et al., 2013). The existence of between-session forgetting in perceptual learning

suggests that repetitive training might not directly lead to the formation of long-term memory; rather, memory after termination of training in each session may decay over time, just like memory in other forms of skill acquisition (Anderson & Hulbert, 2021; Hardt et al., 2013; Jaber, 2005). Many studies have investigated the role of forgetting in stabilizing memory trace via active (Anderson & Hulbert, 2021; Hardt et al., 2013; Storm et al., 2011) and/or passive (Ricker et al., 2016) weakening or elimination of memory traces in the memory domain. Our results could potentially open up a new avenue of research on the role of forgetting in perceptual learning.

Both between-session forgetting and gain in perceptual learning were observed in this study. We found that tasks that required fine feature discrimination near difference thresholds, including vernier-offset discrimination, face-view discrimination, and auditory-frequency discrimination, resulted in between-session forgetting, whereas tasks that used the same stimulus shape and did not require fine feature discrimination, including the contrast detection (only contrast, but not stimulus shape, was varied) and visual shape search tasks, resulted in between-session gain. Consistent with this pattern, results from a large number of studies using the texture discrimination task, which uses constant stimuli, also found between-session gain (Gais et al., 2000; Karni et al., 1994; Stickgold et al., 2000; Tamaki et al., 2020; Yotsumoto et al., 2009). For tasks that use constant stimulus shapes, the same perceptual templates or weight structures are used throughout training (Dosher & Lu, 1999) and could thus be largely retained and might be reactivated during sleep (Gais et al., 2000; Stickgold et al., 2000; Tamaki et al., 2020), resulting in between-session gain. For tasks that are trained near feature-difference threshold, the perceptual templates change with feature-difference threshold during training (e.g., more precise templates are required to discriminate two close orientations) and must be precisely specified to perform near threshold feature discrimination. It might be difficult to remember the details of the most recent perceptual templates in the previous session, resulting in between-session forgetting. This is consistent with the finding that, for an orientation discrimination task, transfer of performance improvement was observed only in low-precision transfer tasks, whereas specificity of performance improvement was observed in high-precision transfer tasks, regardless of the precision of initial training, because it is difficult to carry the details of the perceptual templates required in a high-precision judgment to the transfer task (Jeter et al., 2009). We tentatively conclude that tasks that require precise perceptual templates may exhibit between-session forgetting, and tasks that use constant low-precision perceptual templates may exhibit between-session gain.

In cognitive learning, Rickard (2007) postulated that between-session delay could cause both forgetting and enhanced potential for new learning and therefore rapid relearning. Pinchuk-Yacobi and Sagi (2019) attributed rapid within-session learning to adaptation because of its transient nature. We found that the magnitude of relearning was indeed negatively correlated with between-session effects (off-line gain or forgetting) in the contrast detection, vernier-offset discrimination, face-view discrimination, visual shape search, and auditory-frequency discrimination tasks (see Table S7). On the other hand, there was no significant correlation between within-session adaptation and relearning or between-session gain or forgetting.

To examine the potential influence of training sequence (Kattner et al., 2017) on our results, we fitted the general learning model with two parameters—the best-fitting model of the average data over all participants and the full model with 24 parameters—to the average data in the seven tasks in seven different training sequences. The model-comparison results from different training sequences were, in most cases, consistent with those from fitting the average data across training sequences (see Table S8 and Fig. S5 in the Supplemental Material). In other words, training sequence did not qualitatively impact our results. In addition to training sequence, the results from the new experiment with two training tasks were largely consistent with those within the larger seven-task experiment, demonstrating the robustness of the observed short- and long-term learning components for the same task under different training conditions.

A complete taxonomy of the component processes in perceptual learning will require a more systematic analysis of many different training tasks and conditions. Although the proposed multiprocess model can reasonably explain a variety of learning characteristics across different tasks that varied along multiple dimensions, including sensory modality, stimulus characteristics and structure (Wang et al., 2016), task complexity (Fine & Jacobs, 2002), difficulty (Liu et al., 2012), and training sequence (Kattner et al., 2017), the theoretical framework and its initial applications are the first steps in the process. In the current study, we identified both long- and short-term processes engaged by normal adult participants in seven laboratory tasks. People in other age groups, in other clinical conditions, or who are learning other laboratory or real-world tasks may engage different long- and short-term processes. Therefore, the particular component processes identified in this study may not generalize to them. However, the multicomponent theoretical framework developed in this study might help identify the long- and short-term processes in those situations.

Most previous research analyzed learning curves in session-by-session scale and revealed significant learning improvements across sessions for most tasks, including the seven tasks used in the current study (Amitay et al., 2005; Beard et al., 1995; Bi et al., 2010; Jaeggi et al., 2008; Sigman & Gilbert, 2000). Our study, on the other hand, highlights the importance of fine-grained analysis of the learning curve. In almost all perceptual-learning studies, the learning curve is constructed from scores based on hundreds or even thousands of trials and therefore is very coarse grained. Several recent studies have emphasized the importance of developing trial-by-trial analysis techniques of the learning curve (Kattner et al., 2017; Zhang et al., 2019). In this study, we estimated block-by-block thresholds (96/100 trials) with a maximum likelihood procedure, taking into account information from all trials in each block. Analysis of the learning curves at the block level rather than the typical session level revealed both short- and long-term processes in perceptual learning. Future studies may conduct even more fine-grained analysis of the learning curves at the trial-by-trial level within a hierarchical Bayesian model framework (Zhao et al., 2020).

The proposed multicomponent theoretical framework provides not only a framework to answer one of the fundamental questions in perceptual learning, that is, how perceptual improvements accumulate over time, but also a practical tool to optimize learning outcomes. By manipulating various components of training (e.g., training intensity, allocation of trials, delay between trials and sessions) and studying their effects on the long- and short-term processes in perceptual learning, we could design more efficient training protocols to maximize general learning, within-session relearning, and between-session gain and to minimize within-session adaptation and between-session forgetting for normal as well as clinical populations (Dosher & Lu, 2020).

Transparency

Open Practices

All data have been made publicly available via OSF and can be accessed at https://doi.org/10.17605/OSF.IO/9DBX3. The design and analysis plan for the study were not preregistered. This article has received the badge for Open Data. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

ORCID iD

Zhong-Lin Lu (iD) https://orcid.org/0000-0002-7295-727X

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/09567976211056620

References

Aberg, K. C., Tartaglia, E. M., & Herzog, M. H. (2009). Perceptual learning with chevrons requires a minimal number of trials, transfers to untrained directions, but does not require sleep. *Vision Research*, *49*(16), 2087–2094. https://doi.org/10.1016/j.visres.2009.05.020

Amitay, S., Hawkey, D. J. C., & Moore, D. R. (2005). Auditory frequency discrimination learning is affected by stimulus variability. *Perception & Psychophysics*, *67*(4), 691–698. https://doi.org/10.3758/bf03193525

Anderson, M. C., & Hulbert, J. C. (2021). Active forgetting: Adaptation of memory by prefrontal control. *Annual Review of Psychology*, *72*, 1–36. https://doi.org/10.1146/annurev-psych-072720-094140

Bang, J. W., Shibata, K., Frank, S. M., Walsh, E. G., Greenlee, M. W., Watanabe, T., & Sasaki, Y. (2018). Consolidation and reconsolidation share behavioural and neurochemical mechanisms. *Nature Human Behaviour*, *2*(7), 507–513. https://doi.org/10.1038/s41562-018-0366-8

Beard, B. L., Levi, D. M., & Reich, L. N. (1995). Perceptual learning in parafoveal vision. *Vision Research*, *35*(12), 1679–1690. https://doi.org/10.1016/0042-6989(94)00267-p

Bi, T., Chen, N., Weng, Q., He, D., & Fang, F. (2010). Learning to discriminate face views. *Journal of Neurophysiology*, *104*(6), 3305–3311. https://doi.org/10.1152/jn.00286.2010

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Censor, N., Karni, A., & Sagi, D. (2006). A link between perceptual learning, adaptation and sleep. *Vision Research*, *46*(23), 4071–4074. https://doi.org/10.1016/j.visres.2006.07.022

De Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1–56. https://doi.org/10.1111/0023-8333.00110

Dosher, B. A., & Lu, Z.-L. (1999). Mechanisms of perceptual learning. *Vision Research*, *39*(19), 3197–3221. https://doi.org/10.1016/s0042-6989(99)00059-0

Dosher, B. A., & Lu, Z.-L. (2007). The functional form of performance improvements in perceptual learning: Learning rates and transfer. *Psychological Science*, *18*(6), 531–539. https://doi.org/10.1111/j.1467-9280.2007.01934.x

Dosher, B. A., & Lu, Z.-L. (2020). *Perceptual learning: How experience shapes visual perception*. MIT Press.

Draguhn, A. (2018). Making room for new memories. *Science*, *359*(6383), 1461–1462. https://doi.org/10.1126/science.aat1493

Ebbinghaus, H. (2013). Memory: A contribution to experimental psychology. *Annals of Neurosciences*, *20*(4), 155–156. https://doi.org/10.5214/ans.0972.7531.200408

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fine, I., & Jacobs, R. A. (2002). Comparing perceptual learning across tasks: A review. *Journal of Vision*, *2*(2), Article 5. https://doi.org/10.1167/2.2.5

Frank, S. M., Qi, A., Ravasio, D., Sasaki, Y., Rosen, E. L., & Watanabe, T. (2020). Supervised learning occurs in visual perceptual learning of complex natural images. *Current Biology*, *30*(15), 2995–3000. https://doi.org/10.1016/j.cub.2020.05.050

Gais, S., Plihal, W., Wagner, U., & Born, J. (2000). Early sleep triggers memory for early visual discrimination skills. *Nature Neuroscience*, *3*(12), 1335–1339. https://doi.org/10.1038/81881

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: The role of active forgetting in memory. *Trends in Cognitive Sciences*, *17*(3), 111–120. https://doi.org/10.1016/j.tics.2013.01.001

Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2008). Robust perceptual learning of faces in the absence of sleep. *Vision Research*, *48*(28), 2785–2792. https://doi.org/10.1016/j.visres.2008.09.003

Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2009). How much practice is needed to produce perceptual learning? *Vision Research*, *49*(21), 2624–2634. https://doi.org/10.1016/j.visres.2009.08.022

Jaber, M. (2005). Learning and forgetting models and their applications. In A. B. Badiru (Ed.), *Handbook of industrial and systems engineering* (pp. 30–31). CRC Press.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, USA*, *105*(19), 6829–6833. https://doi.org/10.1073/pnas.0801268105

Jeter, P. E., Dosher, B. A., Petrov, A., & Lu, Z.-L. (2009). Task precision at transfer determines specificity of perceptual learning. *Journal of Vision*, *9*(3), Article 1. https://doi.org/10.1167/9.3.1

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature*, *365*(6443), 250–252. https://doi.org/10.1038/365250a0

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*(5172), 679–682. https://doi.org/10.1126/science.8036518

Kattner, F., Cochrane, A., Cox, C. R., Gorman, T. E., & Green, C. S. (2017). Perceptual learning generalization from sequential perceptual training as a change in learning rate. *Current Biology*, *27*(6), 840–846. https://doi.org/10.1016/j.cub.2017.01.046

Levi, D. M., Polat, U., & Hu, Y.-S. (1997). Improvement in vernier acuity in adults with amblyopia: Practice makes better. *Investigative Ophthalmology & Visual Science*, *38*(8), 1493–1510.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477.

Li, X., Lu, Z.-L., Xu, P., Jin, J., & Zhou, Y. (2003). Generating high gray-level resolution monochrome displays with conventional computer graphics cards and color monitors. *Journal of Neuroscience Methods*, *130*(1), 9–18. https://doi.org/10.1016/s0165-0270(03)00174-2

Liu, J., Lu, Z.-L., & Dosher, B. A. (2012). Mixed training at high and low accuracy levels leads to perceptual learning without feedback. *Vision Research*, *61*, 15–24. https://doi.org/10.1016/j.visres.2011.12.002

Lu, Z.-L., & Dosher, B. (2013). *Visual psychophysics: From laboratory to theory*. MIT Press.

Mascetti, L., Muto, V., Matarazzo, L., Foret, A., Ziegler, E., Albouy, G., Sterpenich, V., Schmidt, C., Degueldre, C., Leclercq, Y., Phillips, C., Luxen, A., Vandewalle, G., Vogels, R., Maquet, P., & Balteau, E. (2013). The impact of visual perceptual learning on sleep and local slow-wave initiation. *The Journal of Neuroscience*, *33*(8), 3323–3331. https://doi.org/10.1523/JNEUROSCI.0763-12.2013

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*(3), 679–690. https://doi.org/10.3758/s13428-010-0049-5

Myung, I. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*, 90–100. https://doi.org/10.1016/S0022-2496(02)00028-7

Neath, A. A., & Cavanaugh, J. E. (2012). The Bayesian information criterion: Background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*(2), 199–203.

Pavlik, P. I., Jr., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, *29*(4), 559–586. https://doi.org/10.1207/s15516709cog0000_14

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Pinchuk-Yacobi, N., & Sagi, D. (2019). Orientation-selective adaptation improves perceptual grouping. *Journal of Vision*, *19*(9), Article 6. https://doi.org/10.1167/19.9.6

Rickard, T. C. (2007). Forgetting and learning potentiation: Dual consequences of between-session delays in cognitive skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(2), 297–304. https://doi.org/10.1037/0278-7393.33.2.297

Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *Quarterly Journal of Experimental Psychology*, *69*(10), 1969–1995. https://doi.org/10.1080/17470218.2014.914546

Sagi, D. (2011). Perceptual learning in vision research. *Vision Research*, *51*(13), 1552–1566. https://doi.org/10.1016/j.visres.2010.10.019

Shibata, K., Sasaki, Y., Bang, J. W., Walsh, E. G., Machizawa, M. G., Tamaki, M., Chang, L.-H., & Watanabe, T. (2017). Overlearning hyperstabilizes a skill by rapidly making neurochemical processing inhibitory-dominant. *Nature Neuroscience*, *20*(3), 470–475. https://doi.org/10.1038/nn.4490

Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neuroscience*, *3*(3), 264–269.

Stickgold, R., James, L., & Hobson, J. A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience*, *3*(12), 1237–1238. https://doi.org/10.1038/81756

Storm, B., Angello, G., & Bjork, E. (2011). Thinking can cause forgetting: Memory dynamics in creative problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1287–1293. https://doi.org/10.1037/a0023921

Tamaki, M., Berard, A. V., Barnes-Diana, T., Siegel, J., Watanabe, T., & Sasaki, Y. (2020). Reward does not facilitate visual perceptual learning until sleep occurs. *Proceedings of the National Academy of Sciences, USA*, *117*(2), 959–968. https://doi.org/10.1073/pnas.1913079117

Walker, M., Brakefield, T., Hobson, J., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, *425*, 616–620. https://doi.org/10.1038/nature01930

Wang, R., Wang, J., Zhang, J.-Y., Xie, X.-Y., Yang, Y.-X., Luo, S.-H., Yu, C., & Li, W. (2016). Perceptual learning at a conceptual level. *The Journal of Neuroscience*, *36*(7), 2238–2246. https://doi.org/10.1523/jneurosci.2732-15.2016

Watanabe, T., & Sasaki, Y. (2015). Perceptual learning: Toward a comprehensive theory. *Annual Review of Psychology*, *66*(1), 197–221. https://doi.org/10.1146/annurev-psych-010814-015214

Yang, J., Yan, F.-F., Chen, L., Xi, J., Fan, S., Zhang, P., Lu, Z.-L., & Huang, C.-B. (2020). General learning ability in perceptual learning. *Proceedings of the National Academy of Sciences, USA*, *117*(32), 19092–19100. https://doi.org/10.1073/pnas.2002903117

Yotsumoto, Y., Sasaki, Y., Chan, P., Vasios, C. E., Bonmassar, G., Ito, N., Náñez, J. E., Shimojo, S., & Watanabe, T. (2009). Location-specific cortical activation changes during sleep after training for perceptual learning. *Current Biology*, *19*(15), 1278–1282. https://doi.org/10.1016/j.cub.2009.06.011

Zhang, P., Zhao, Y., Dosher, B. A., & Lu, Z.-L. (2019). Assessing the detailed time course of perceptual sensitivity change in perceptual learning. *Journal of Vision*, *19*(5), Article 9. https://doi.org/10.1167/19.5.9

Zhao, Y., Zhang, P., Chen, G., Yang, J., Huang, C.-B., Liu, J., Dosher, B. A., & Lu, Z.-L. (2020). Evaluating the functional form of perceptual learning with trial-by-trial analysis. *Journal of Vision*, *20*(11), Article 1643. https://doi.org/10.1167/jov.20.11.1643