Perception & Psychophysics
2003, 65 (4), 632-648

Sequential priming is not constrained by the shape of long-term learning curves

SATORU SUZUKI and BRIAN A. GOOLSBY
Northwestern University, Evanston, Illinois

When multiple stimulus-to-response (S–R) mappings are randomly intermixed and repeated in a block of trials, immediate repetitions of an aspect of a stimulus and/or a response can facilitate stimulus detection, classification, and/or response selection—known as *sequential priming*. In addition to these short-term effects, response times (RTs) for almost any task diminish with extended practice; improvements can occur over many days, and RT learning curves typically assume exponential or power functions. We investigated whether short-term sequential priming and long-term practice modulate RT through a common mechanism, using a variant of the additive factors method. We tracked how various priming effects, presumably affecting different processing stages (e.g., stimulus selection, stimulus identification/classification, and S–R mapping), varied over training sessions as RT diminished. All the priming effects either were not reduced or reduced approximately linearly at rates much slower than those predicted by the shapes of the corresponding RT learning curves. The overall results suggest that short-term sequential priming and long-term practice modulate RT through relatively separate mechanisms, even though they appear to affect a common set of behaviorally defined processing stages.

Performance in almost any task (be it perceptual, cognitive, or motor) improves with long-term practice. For example, when a target must be selected among distractors, as in visual search tasks, practice may increase the salience of a target against distractors and/or facilitate direction of attention to a target through a variety of mechanisms, such as facilitation of relevant feature detectors, efficient selection of diagnostic target features, perceptual unitization of complex search items, utilization of specific memory traces from prior search, increasingly global deployment of attention, and/or efficient grouping or suppression of distractors (e.g., Ahissar & Hochstein, 1993, 1996; Ahissar, Laiwand, Kozminsky, & Hochstein, 1998; Czerwinski, Lightfoot, & Shiffrin, 1992; Hillstrom & Logan, 1998; Shiffrin & Lightfoot, 1997; Shiffrin & Schneider, 1977; Sireteanu & Rettenbach, 2000; Treisman, Vieira, & Hays, 1992). When target selection and identification are relatively trivial (e.g., identifying a target presented with no distractors by its position, color, or familiar name) and the task requires only simple stimulus-to-response (S–R) mappings (e.g., only one target stimulus is mapped to each response finger), practice effects may be primarily due to a strengthening of the few S–R connections that are used consistently and repeatedly (e.g., Kirby, 1980; Soetens, Deboeck, & Hueting, 1984; Welford, 1980). When a task requires formations of complex S–R mappings, such as when an arbitrary, uncategorizable set of target symbols are mapped to each response finger, practice effects may be mediated primarily by developing or choosing an increasingly efficient cognitive strategy (e.g., attending to a more relevant set of visual features) and/or by forming increasingly task-optimized categorical representations (e.g., Newell & Rosenbloom, 1981; Pashler & Baylis, 1991a, Experiments 2 and 3). In addition to improvements in processing strategies, representations, and/or S–R mappings, practice effects may also be mediated by the storing of each instance of making a particular S–R (or stimulus-interpretation-to-response) connection and the utilization of the stored information in future trials, thus bypassing relatively inefficient general purpose S–R processing algorithms (e.g., Logan, 1988, 1990; Treisman et al., 1992). Finally, when a task requires fine image processing, such as discriminations among very similar patterns and detection or identification of stimuli under degraded conditions (e.g., low contrast, brief presentation, added noise, and masking), practice effects seem to be mediated primarily by modifications of specific perceptual mechanisms that selectively process the relevant visual features (e.g., Ball & Sekuler, 1982; Fahle, 1994; Fiorentini & Berardi, 1980; Polat & Sagi, 1994). Even this brief survey indicates that practice optimizes a variety of processing stages, including stimulus selection, feature detection, identification, classification, and S–R mapping.

In addition to long-term incremental facilitation of performance, performing a particular S–R trial can pro-

---

A part of this work was presented at the 41st Annual Meeting of the Psychonomic Society (November 2000). This work was supported by National Science Foundation Grant SBR-9817643 given to the first author. We thank Hal Pashler, Robert Proctor, Kimron Shapiro, and Eric Soetens for their constructive and helpful comments. We also thank Marcia Grabowecky and Sania Hamilton for helping to make the manuscript clearer. Correspondence concerning this article should be addressed to S. Suzuki, Department of Psychology, Northwestern University, 2029 Sheridan Rd., Evanston, IL 60208 (e-mail: satoru@northwestern.edu).

duce a short-lasting priming effect on subsequent trials. In general, when a given S–R trial is immediately preceded by an identical S–R trial, response time (RT) is substantially reduced relative to when it is immediately preceded by a different S–R trial (first-order priming effects). These automatic sequential priming effects are evident when S–R mappings are sufficiently complex (e.g., arbitrary mapping, such as shape-to-finger mapping, rather than compatible mapping, such as left-position-to-left-finger and right-position-to-right-finger mapping) and when response-to-stimulus intervals (RSIs) are sufficiently short; when S–R mappings are simple and RSIs are relatively long, *expectancy* effects may also influence RT (see Kirby, 1976, 1980; Soetens, 1998; Soetens, Boer, & Hueting, 1985; and Soetens, Deboeck, & Hueting, 1984, for details about how automatic priming and expectancy influence RT in serial choice RT tasks). Here, we focus on automatic sequential priming effects (potential contributions of expectancy in our tasks will be discussed later, in the Results section).

Sequential priming is evident when different S–R trials are randomly intermixed and repeatedly presented in a block of trials. Since each S–R trial is frequently repeated, episodic memory and long-term cumulative effects are expected to contribute minimally to sequential priming effects; on any given trial, an observer would have encountered each type of S–R trial numerous times in the recent past. Sequential priming is thus considered to reflect a short-lasting automatic potentiation of perceptual–response pathways (or channels) that have just been used in a recent trial, automatic inhibition of competing pathways, or both (e.g., Soetens, Deboeck, & Hueting, 1984; Vervaeck & Boer, 1980). Sequential priming effects have been demonstrated in visual search tasks (e.g., Goolsby & Suzuki, 2001; Hillstrom, 2000; Maljkovic & Nakayama, 1994, 2000; Rabbitt, Cumming, & Vyas, 1979), as well as in more traditional stimulus classification tasks (e.g., Campbell & Proctor, 1993; Pashler & Baylis, 1991b; Soetens, 1998; see Kirby, 1980, and Kornblum, 1973, for early reviews).

Although there have been numerous studies of long-term practice effects and sequential priming effects on RT, there have been few studies of the potential relationship between them. For example, it is possible that effects of long-term practice result from a temporal accumulation of small residual priming effects, assuming that priming has a decay function with a long tail (e.g., Campbell & Proctor, 1993; Pashler & Baylis, 1991b). Alternatively, long-term practice and sequential priming might modulate different processing stages or operate through separate mechanisms. By comparing how long-term practice and first-order sequential priming transferred from one condition to another (see the General Discussion section for details), Pashler and Baylis (1991a, 1991b) and Campbell and Proctor (1993) suggested that long-term practice and sequential priming both facilitated a relatively high-level stimulus-category-to-response mapping, whereas sequential priming additionally facilitated a relatively low-level S–R mapping (e.g., Pashler & Baylis, 1991b).

In the present study, we examined whether or not long-term practice and sequential priming operated through a common mechanism. Even if practice and priming affected the same set of behaviorally defined processing stages, they might influence those stages through separate mechanisms. A well-known method for assessing whether two factors have separate (noninteractive) influences on behavior is to assess whether the effects of the two factors are additive (additive factors methods; e.g., Sternberg, 1969). In the simplest case, if practice effects and sequential priming effects are numerically additive, the amount of priming should remain constant over practice sessions; in other words, short-term RT reductions due to repetitions of stimuli, responses, or both (relative to nonrepeated cases) should remain constant regardless of the changes in overall RT due to practice. Such a result would suggest that long-term practice and sequential priming have noninteractive influences on RT, likely through parallel mechanisms.

Previous studies have demonstrated that sequential priming for simple S–R tasks (e.g., press left key for left target and right key for right target) decreased with extended practice (e.g., Kirby, 1980; Soetens et al., 1985; Soetens, Deboeck, & Hueting, 1984; Vervaeck & Boer, 1980). However, in these studies, it was not examined in detail how reductions in priming were related to decreases in overall RT. Even if priming is reduced over practice sessions, this does not necessarily imply that practice and priming influence RT through a common mechanism. We will argue that if practice and priming do operate through a common mechanism, practice-induced changes in priming should be closely related to the shape of the corresponding RT learning curve.

Practice-induced improvements are usually initially rapid, but the gain gradually diminishes over time until performance reaches an asymptote. For a variety of tasks, learning curves (performance improvements as a function of the amount of practice) can be well fit by power functions of the form [RT or Error] = $a + b(X+c)^{-d}$ or by exponential functions of the form [RT or Error] = $f + g \cdot e^{-\alpha X}$, where X indicates the amount of practice (typically the number of practice sessions) and a, b, c, d, f, g, and α are positive constants; a and f represent asymptotic RTs (or errors), b and g represent the overall RT (or error) reduction through practice, c represents the degree of prior practice, and d ($d/[X+c]$, to be exact) and α represent rates of improvement (in terms of proportions by which RT approaches asymptotes). Which of the two mathematical forms represents learning curves more accurately is currently being debated (e.g., Heathcote, Brown, & Mewhort, 2000; Newell & Rosenbloom, 1981).

Nevertheless, regardless of their exact mathematical forms, RT learning curves demonstrate examples of the law of diminishing returns. A given amount of practice input (e.g., performing a block of practice trials) initially produces a large reduction in RT, but the same amount of practice input produces a smaller reduction in RT as learning progresses. In other words, practice effects operate through a mechanism that scales each practice input by the diminishing rate of learning—that is, by the first derivative of the RT learning curve, dY_{learn}/dX. If sequential priming also affected RT through the same mechanism, each priming input (e.g., a prior presentation of the current stimulus in the immediately preceding trial) should also be scaled by dY_{learn}/dX when priming is measured as changes in RT. Thus, *if long-term practice and sequential priming facilitated a common processing stage through the same mechanism, priming effects should be scaled by* dY_{learn}/dX *over practice sessions, just as practice effects are*. In other words, priming should decrease over practice sessions, and the decrease should be well fit by $C\cdot[dY_{\text{learn}}/dX]$, where C is the fitting parameter that discounts the overall scale difference between practice and priming effects. As will be discussed below, however, this argument is complicated by the fact that a single RT learning curve is likely to reflect concurrent optimization of multiple underlying processing stages.

Intuitively, any simple S–R task requires at least the following three stages of processing: (1) stimulus selection (finding the stimulus to be processed), (2) stimulus identification and categorization, and (3) response selection; note that these stages may or may not be executed sequentially. It is reasonable to assume that the observed learning curve represents a linear combination of the component learning curves that represent practice effects on these and other processing stages (possibilities of nonlinear contributions will be considered in the General Discussion section). Then, it is straightforward to show that the observed learning curve approaches its asymptote at a faster rate (in terms of proportions by which RT approaches asymptotes) than does the slowest component learning curve and at a slower rate than does the fastest component learning curve (see the Appendix). More relevant to the following discussion, this means that *if different component learning curves approached their asymptotes at different rates, some of the component learning curves should approach their asymptotes faster than does the observed learning curve, whereas other component learning curves should approach their asymptotes slower than does the observed learning curve*.

We used tasks that allowed us to measure multiple priming effects, presumably affecting each of the three major processing stages mentioned above (stimulus selection, stimulus identification/categorization, and response selection). As observers underwent extended practice sessions, we monitored changes in priming effects, as well as reductions in overall RT (learning curves). The following outcomes were expected.

First, if practice predominantly optimized one processing stage (e.g., stimulus selection) through a mechanism that scaled inputs by dY_{learn}/dX, the corresponding priming effect (e.g., priming for stimulus selection) should also be scaled by dY_{learn}/dX, if practice and priming operated through the same mechanism. Similarly, if practice optimized all processing stages at an equal rate, all priming effects should be scaled by dY_{learn}/dX.

Alternatively, if practice optimized all or some of the processing stages at different rates, component learning curves for some stages should have faster learning rates than does the overall learning curve, whereas component learning curves for other stages should have slower learning rates than does the overall learning curve (when learning rates are measured as proportions by which RT approaches asymptotes). Thus, so long as the measured priming effects reflected all of the stages affected by practice, some of the priming effects also should diminish faster than would be predicted by dY_{learn}/dX, whereas other priming effects should diminish slower than would be predicted by dY_{learn}/dX, if practice and priming operated through the same mechanism.

In contrast, if all the measured priming effects diminished slower than would be predicted by dY_{learn}/dX (or did not diminish), we would have two possible interpretations. First, practice happened to primarily affect a processing stage untapped by any of the measured priming effects; if priming could be measured for that processing stage, it would diminish according to dY_{learn}/dX over practice sessions. Second, practice and priming might operate through different mechanisms even if they affect common processing stages. For example, a given processing stage, such as stimulus selection, might be facilitated separately by a long-term mechanism and a short-term mechanism. To decide between these two potential interpretations, it was important to use tasks that involved priming manipulations affecting multiple processing stages. Inclusion of the stimulus selection stage was particularly relevant because studies on visual search have shown that both long-term practice (e.g., Ahissar & Hochstein, 1993, 1996; Ahissar et al., 1998; Czerwinski et al., 1992; Hillstrom & Logan, 1998; Shiffrin & Lightfoot, 1997; Shiffrin & Schneider, 1977; Sireteanu & Rettenbach, 2000; Treisman et al., 1992) and sequential priming (e.g., Goolsby & Suzuki, 2001; Maljkovic & Nakayama, 1994, 2000) facilitate stimulus selection and that priming of stimulus selection can be functionally dissociated from priming of stimulus identification/categorization and response selection (Goolsby & Suzuki, 2001). We thus adopted the basic paradigm developed by Maljkovic and Nakayama (1994, 2000; see also Goolsby & Suzuki, 2001), which allowed us to monitor stimulus selection priming, as well as priming related to stimulus identification and response.

In each display, three diamond shapes were presented. Each diamond had its right or left side cut off (chipped). Targets were defined by odd color, either a green diamond presented among red distractor diamonds or a red diamond presented among green distractor diamonds (Figure 1). In each trial, an observer found the odd-colored diamond and reported which side of it was chipped, by pressing the left key when the left side was chipped and pressing the right key when the right side was chipped—called the *chip task*. By using this paradigm, two types of

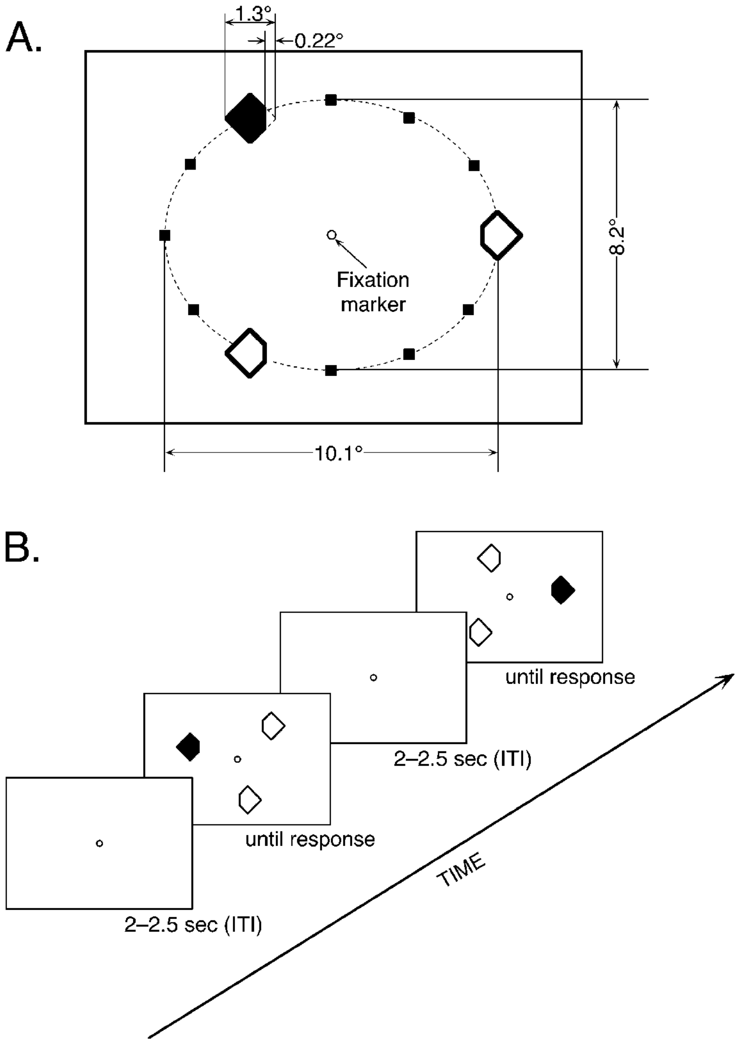

Figure 1. Stimuli and procedure. (A) The 12 possible stimulus locations are shown with an example of a target (the black figure) and two distractors (white figures). To keep visual acuity approximately equal for each stimulus location, the stimuli were placed on an invisible ellipse. The target did not appear at locations directly above or below the fixation marker in the hemifield and the hemifield–chip tasks. The two distractors were always separated by four location steps (or 120° rotation) from the target and from each other. (B) A trial sequence is shown schematically for two consecutive trials. The interval between successive stimulus displays (intertrial interval, or ITI) was randomly varied between 2 and 2.5 sec to prevent observers from anticipating the onset of the stimulus display.

sequential priming effects were measured: (1) color combination priming and (2) stimulus (or response) priming. The RT for a given trial was facilitated when the same color combination had been encountered in the previous trial. When the color combination was switched randomly between green-target-among-red-distractors and red-target-among-green-distractors across trials, the priming effect due to a single trial lasted through up to six trials; that is, we obtained significant benefits of having had the same color combination six trials back while colors changed randomly in the intervening trials. We previously had shown that this priming speeded target detection by facilitating the direction of attention to the odd-colored target (Goolsby & Suzuki, 2001), demonstrating that color combination priming is a priming of stimulus selection. Although not as robust as the color combination (stimulus selection) priming, the RT for a given trial was also speeded when the same side of the target diamond had been chipped in a previous trial. Since both the side of chip and the response finger were repeated in this case, the facilitation could be due to shape repetition, response repetition, or both.

In the present study, we also added two new variants of this task. To widen the potential generality of the results, we introduced two levels of task difficulty and measured a larger variety of priming effects. The stimuli themselves were identical to those described above. In the easy task, observers reported whether the odd-colored target was in the left or the right visual hemifield (the *hemifield task*). This task was considerably easier than the high-acuity task of determining which side of the target diamond was chipped. For the hemifield task, we measured priming effects that were due to repetitions of color combination and those that were due to repetitions of the hemifield in which the target was presented (confounded with response repetition).

In the difficult task, a more complex S–R mapping was used. An observer reported whether the side of chip and the visual hemifield location of the target matched (i.e., a *left*-side-chipped target presented in the *left* visual hemifield or a *right*-side-chipped target presented in the *right* visual hemifield) or not. We called this task the *hemifield–chip task*. This task enabled us to monitor multiple priming effects: those due to repetitions of color combination, repetitions of hemifield location, repetitions of target shape (i.e., repetitions of the side of chip), and repetitions of response (unconfounded by other factors).

We were thus able to measure priming that affected (1) stimulus selection (color combination priming in all three tasks), (2) stimulus property identification/classification (hemifield priming and side-of-chip priming in the hemifield–chip task), and (3) S–R mapping or response selection (side-of-chip/response priming in the chip task, hemifield/response priming in the hemifield task, and response priming in the hemifield–chip task). We measured individual contributions of these priming effects by varying color combination, hemifield location, side of chip, and/or response finger randomly from trial to trial.

For our purposes, this paradigm provided some advantages over the more traditional information reduction paradigm (IRP; see Bertelson, 1965, Rabbitt, 1968, and Smith, 1968, for early references). In the IRP, the relative contributions of response repetition and stimulus repetition are assessed by assigning multiple stimuli to each response. Response repetition priming can be assessed by examining the effects of repeating or not repeating a response when the stimulus changes. Stimulus repetition priming (over and above response repetition priming) can be assessed by examining the effects of repeating or not repeating an identical stimulus when the response remains the same. One potential disadvantage of the IRP is that stimulus repetition is always confounded by response repetition—that is, a stimulus never repeats without also repeating a response. This potentially raises problems when the individual contributions of response priming and stimulus priming need to be assessed. Since stimulus priming is always measured *above* response priming, a large benefit from response priming could potentially reduce stimulus priming, due to floor effects. In the present paradigm, stimulus features and responses were varied randomly (except for the cases in which they were confounded). For example, a color combination was repeated equally often regardless of whether a response was repeated or not. A second advantage (as was stressed earlier) is that the present paradigm allowed us to monitor all of the potentially critical stages of performing an S–R task: stimulus selection, stimulus identification/categorization, and response selection. In particular, none of the prior studies that examined effects of practice on sequential priming included priming of stimulus selection.

METHOD

Overview

Observers performed a block of 200 trials (per task) in each practice session. Practice sessions continued for days (up to months) in order to obtain well-defined learning curves. The goal was to compare practice-induced changes in various priming effects with the shape of the RT learning curve. In particular, we examined whether any of the priming effects diminished in proportion to the first derivative of the learning curve, dY_{learn}/dX.

Because it is a matter of debate as to whether power functions or exponential functions more appropriately represent learning curves (e.g., Heathcote et al., 2000; Newell & Rosenbloom, 1981), we used both functions to fit the learning curves—$Y(X) = a + b(X + c)^{-d}$ (power) and $Y(X) = f + g \cdot e^{-\alpha X}$ (exponential). Both functions fit our data well and produced indistinguishable results; however, note that, for our purposes, the exact mathematical form is irrelevant so long as the functions provide appropriate fits to the learning curves. The first derivatives took the forms $C \cdot (X + c_0)^{-(d_0+1)}$ (power) and $C \cdot e^{-\alpha_0 X}$ (exponential), where c_0, d_0, and α_0 were determined by fitting the RT learning curves. Because the relative strengths of practice and priming effects were unknown, the scaling factor C was used as the fitting parameter to fit changes in priming effects over practice sessions. To reiterate the predictions in more quantitative terms, we expected the following outcomes.

First, if practice predominantly optimized one processing stage through a mechanism that scaled inputs by dY_{learn}/dX, the corresponding priming effect should be well fit by $C \cdot [dY_{\text{learn}}/dX]$—that is, by $C \cdot (X + c_0)^{-(d_0+1)}$ and $C \cdot e^{-\alpha_0 X}$—if practice and priming operated through the same mechanism. Similarly, if practice optimized all processing stages at an equal rate, all priming effects should be well fit by $C \cdot (X + c_0)^{-(d_0+1)}$ and $C \cdot e^{-\alpha_0 X}$.

Alternatively, if practice optimized all or some of the processing stages at different rates, component learning curves for some stages should have ds and αs equal to or greater than d_0 and α_0 (obtained for the observed RT learning curve), whereas component learning curves for other processing stages should have ds and αs equal to or smaller than d_0 and α_0 (see the Appendix). Note that the relevant learning rates beyond scaling factors (i.e., proportional learning rates, as defined in the Appendix) are represented by d and α for power and exponential functions, respectively. Thus, so long as the measured priming effects spanned all of the stages affected by practice, some of the priming effects should diminish more quickly than $C \cdot (X + c_0)^{-(d_0+1)}$ and $C \cdot e^{-\alpha_0 X}$ and, thus, be better fit with $(d+1)$ and α larger than (d_0+1) and α_0, whereas other priming effects should diminish more slowly than $C \cdot (X + c_0)^{-(d_0+1)}$ and $C \cdot e^{-\alpha_0 X}$ and, thus, be better fit with $(d+1)$ and α smaller than (d_0+1) and α_0, if practice and priming operated through the same mechanism.

If all the measured priming effects diminished more slowly than $C \cdot (X + c_0)^{-(d_0+1)}$ and $C \cdot e^{-\alpha_0 X}$, we would have two possible interpretations. First, practice could have primarily affected a processing stage untapped by any of the measured priming effects. This interpretation, however, would be unlikely, because we monitored multiple priming effects spanning the three major processing stages required for performing any typical S–R task—stimulus selection, stimulus identification/categorization, and response selection—which are also known to be facilitated by long-term practice (see the introduction). Thus, a more likely interpretation would be that practice and priming may affect common processing stages but that they operate through different mechanisms; that is, practice, but not priming, operates through a mechanism that scales inputs by dY_{learn}/dX.

Observers

Three observers—K.O. (a naive and novice observer), S.S. (one of the authors), and Y.S. (a naive observer with experience as a participant in psychophysical tasks)—were trained extensively in the hemifield task and the hemifield–chip task. We also analyzed data from 8 observers (naive observers with mixed degrees of experience in perceptual tasks) who performed multiple sessions of the chip task; some aspects of their data had previously been reported in Goolsby and Suzuki (2001). All the observers were tested individually in a dimly lit room.

Apparatus

The stimuli were displayed on a 17-in. color monitor (75 Hz), and all the experiments were controlled with an Apple PowerPC 8600 (300 MHz) computer using the Vision Shell software (micro ML, Inc.).

Stimuli

The standard display consisted of three diamond shapes arranged on an invisible ellipse[1] centered at the fixation marker (Figure 1A). The horizontal axis of the ellipse subtended 10.1° of visual angle, and the vertical axis subtended 8.2°. The diamond shapes could be presented at any of the twelve possible locations along the circumference of this invisible ellipse (at 0°, 30°, 60°, 90°, 120°, . . . , 300°, and 330° locations beginning at due north), with the constraint that the three diamonds were approximately equidistant from one another—that is, separated from one another by four location steps (or by 120° rotation). Each diamond subtended 1.3° × 1.3° visual angle and had its left or right side "chipped," with the depth of chip subtending 0.22°. The side of the chip for each diamond was randomly determined in each trial.

In each stimulus display, one of the three diamonds had a different color than the other two diamonds—a red diamond among two green diamonds or a green diamond among two red diamonds. The color (red or green) of the odd-colored target diamond was determined randomly for each trial. The location of the target was also determined randomly for each trial, but the locations directly above and below the fixation marker were excluded from the hemifield task and the hemifield–chip task, because these locations could not be classified as belonging to the left or the right visual field. The red, CIE[.629, .346], and the green CIE[.299, .598], used were set to be approximately equiluminant. The red had a fixed luminance of 5.0 cd/m^2; the luminance for the green was determined for each observer, using flicker photometry at a frequency of 20 Hz. The stimuli were presented against a dark background (all color guns of the monitor turned off).

The fixation marker, presented at the center of the screen, was a small achromatic open circle (39 cd/m^2, CIE[.262, .282], diameter = 0.26), drawn with a 1-pixel-thick line (each pixel subtending 0.043°). The fixation marker remained on throughout each trial. The observers were seated in a hard-backed chair at 50 cm from the screen. Although head position was not fixed with a chinrest, the observers measured and adjusted their viewing distance before the start of each block of 200 trials.

Procedure

Each trial began with a blank fixation screen presented for an interval of 2–2.5 sec, followed by a stimulus display (color-singleton display) which consisted of the three diamond shapes, one of which was oddly colored. The odd-colored item was the target item for which the observers made a response. The stimulus display remained on until a response had been made. As was mentioned above, there were three tasks: the *hemifield task*, the *hemifield–chip task*, and the *chip task*.

In the hemifield task, the observers reported the visual hemifield (left or right) in which the target item was presented. When the target item was presented in the left visual field, they pressed the "z" key (located on the lower left side of keyboard) with the left index finger, and when the target item was presented in the right visual field, they pressed the "/" key (located on the lower right side of keyboard) with the right index finger.

In the hemifield–chip task, observers responded on the basis of whether or not the side of chip and the hemifield location of the target were compatible. They pressed the up-arrow key with the right index finger when the side of the chip and the hemifield location of the target were the same (i.e., a left-side-chipped target presented in the left visual field or a right-side-chipped target presented in the right visual field) and pressed the down-arrow key with the left index finger when the side of chip and the hemifield location of the target were opposite (i.e., a left-side-chipped target presented in the right visual field or a right-side-chipped target presented in the left visual field). The up-arrow key was located physically above the down-arrow key (i.e., located farther away from the observer than was the down-arrow key on the computer keyboard); this response key arrangement was used, instead of using keys located on the left and right sides of keyboard, in order to reduce the potential effect of the response keys' corresponding or not with the side of chip or the hemifield location of the item (see, e.g., Simon & Wolf, 1963; Soetens, Deboeck, Hueting, & Merckx, 1984).

In the chip task (data from Goolsby & Suzuki, 2001), the observers reported the side of chip of the target. When the left side of the target was chipped, they pressed the "z" key with the left index finger, and when the right side of the target was chipped, they pressed the "/" key with the right index finger. For reasons unrelated to the present goal, for half of the 8 observers, the target location was precued in half of the trials, and for the remaining 4 observers, a warning circle was flashed centrally prior to the stimulus display in half of the trials. The precued trials and the prewarned trials were removed from the analysis here (both as priming trials and as primed trials), although we found that the precued, the prewarned, and the noncued (regular) trials all produced nearly equivalent priming on subsequent noncued trials (Goolsby & Suzuki, 2001).

The observers were instructed to respond as quickly as possible while preserving accuracy above 95%. The response time and response key were recorded on each trial. No feedback was given following incorrect responses. The observers were instructed to maintain eye fixation on the fixation marker prior to the presentation of each stimulus array. Eye movements to the target were allowed after the onset of the stimulus array, as in our previous study (Goolsby & Suzuki, 2001).

In each session (one session per day), the 3 observers (K.O., S.S., and Y.S.) performed a block (200 trials) of the hemifield task and a block of the hemifield–chip task. The order in which the two tasks were performed (hemifield task first or hemifield–chip task first) was alternated across sessions. Observer K.O. participated in 28 sessions over 67 days; S.S. participated in 71 sessions over 124 days; Y.S. participated in 33 sessions over 42 days. Although the observers were encouraged to run their sessions on a daily basis, the intervals between sessions varied somewhat across observers, due to scheduling difficulties. Observer S.S. was run in about twice as many sessions, partly because his learning curves approached apparent asymptotes more gradually than did those for the other observers and partly because he was available to be tested in extended sessions. The 8 observers from Goolsby and Suzuki (2001) performed their first 12 blocks of the chip task (for which data are analyzed) over 3–13 days. The mean data (rather than individual observers' data) were analyzed in this case because fewer trials were available per block to compute priming effects (due to removal of precued and prewarned trials). All the observers initially received a block of 200 training trials (per task) before the practice sessions began.

RESULTS

Long-Term Practice Effects (RT Learning Curves)

We evaluated RTs only for correct responses. The error rates were low; mean error rates were 2.2% for the hemifield task, 3.7% for the hemifield–chip task, and 3.2% for the chip task. Median RT was computed for each practice session for each observer. Individual observers' data were analyzed separately for the hemifield task and the hemifield–chip task; mean data ($N = 8$) were analyzed for the chip task (Table 1).

We note that, for simplicity, the plots shown in Figure 2 are all averaged across observers. This means that, for the hemifield task and the hemifield–chip task, the data from only the first 28 practice sessions are presented. Thus, for these tasks (Figures 2A–2H), the plots are shown primarily to illustrate the general appearances of the learning curves (with exponential-function–based fits shown as an example) and the linear and learning-curve–based fits of the priming data. Tables 1–6 show the analyses of these fits and other effects, using data from all the sessions and using both power and exponential functions.

As is shown in Figure 2 (RT panels: A, D, and I), typical RT learning curves were obtained for all three tasks. As is shown in Table 1A, all the learning curves obtained were well fit by both power functions (explaining about 82%–97% of the variance) and exponential functions (explaining about 81%–97% of the variance).

Sequential Priming Effects

For each block of trials (per practice session per task), we computed various first-order priming effects on correct trials (where both the current trial and the preceding trial were correct). For example, the stimulus selection priming due to repetitions of the same color combination was computed by subtracting the median RT for all the trials for which the color combination (red target among green distractors or green target among red distractors) had been the same in the immediately preceding trial from the median RT for all trials for which the color combination had been different in the immediately preceding trial. *Thus, a positive value indicates the amount by which RT was speeded when a color combination had been the same in the immediately preceding trial relative to when it was different.* The color combination (stimulus selection) priming was evaluated for all three tasks (Figures 2B, 2E, and 2J; Tables 2A, 3A, and 4A). Other priming effects were also evaluated similarly. For the hemifield task, we evaluated the hemifield/response priming (the effect of whether the visual hemifields in which the target was presented and the response finger were the same or different in the immediately preceding trial; Figure 2C; Table 2B). For the hemifield–chip task, we evaluated the hemifield priming (the effect of whether the visual hemifield in which the target was presented was the same or different in the immediately preceding trial; Figure 2F; Table 3B), the shape (side-of-chip) priming (the effect of whether the side of chip of the target was the same or different in the immediately preceding trial; Figure 2G; Table 3C), and the response priming (the effect of whether the response finger was the same or different in the immediately preceding trial; Figure 2H; Table 3D). For the chip task, we evaluated the shape (side-of-chip)/response priming (the effect of whether the side of chip of the target and the response finger were the same or different in the immediately preceding trial; Figure 2K; Table 4B).

Overall priming (averaged across practice sessions) was computed for each priming effect (see under "Overall Priming" in Tables 2–4). The statistical significance was evaluated using practice session as the random effect for the hemifield task and the hemifield–chip task, for which

Table 1
Fitting Response Time Over Sessions

	Power Fit		Exponential Fit		
Observer	d_0	Proportion of Variance Explained	α_0	Proportion of Variance Explained	No. of Sessions
		Hemifield Task			
K.O.	1.497	.816	0.560	.816	28
S.S.	0.693	.923	0.096	.932	71
Y.S.	1.383	.823	0.311	.808	33
		Hemifield–Chip Task			
K.O.	1.090	.966	0.498	.959	28
S.S.	0.264	.945	0.096	.905	71
Y.S.	1.054	.926	0.243	.899	33
		Chip Task			
Mean of 8	0.083	.958	0.263	.968	12

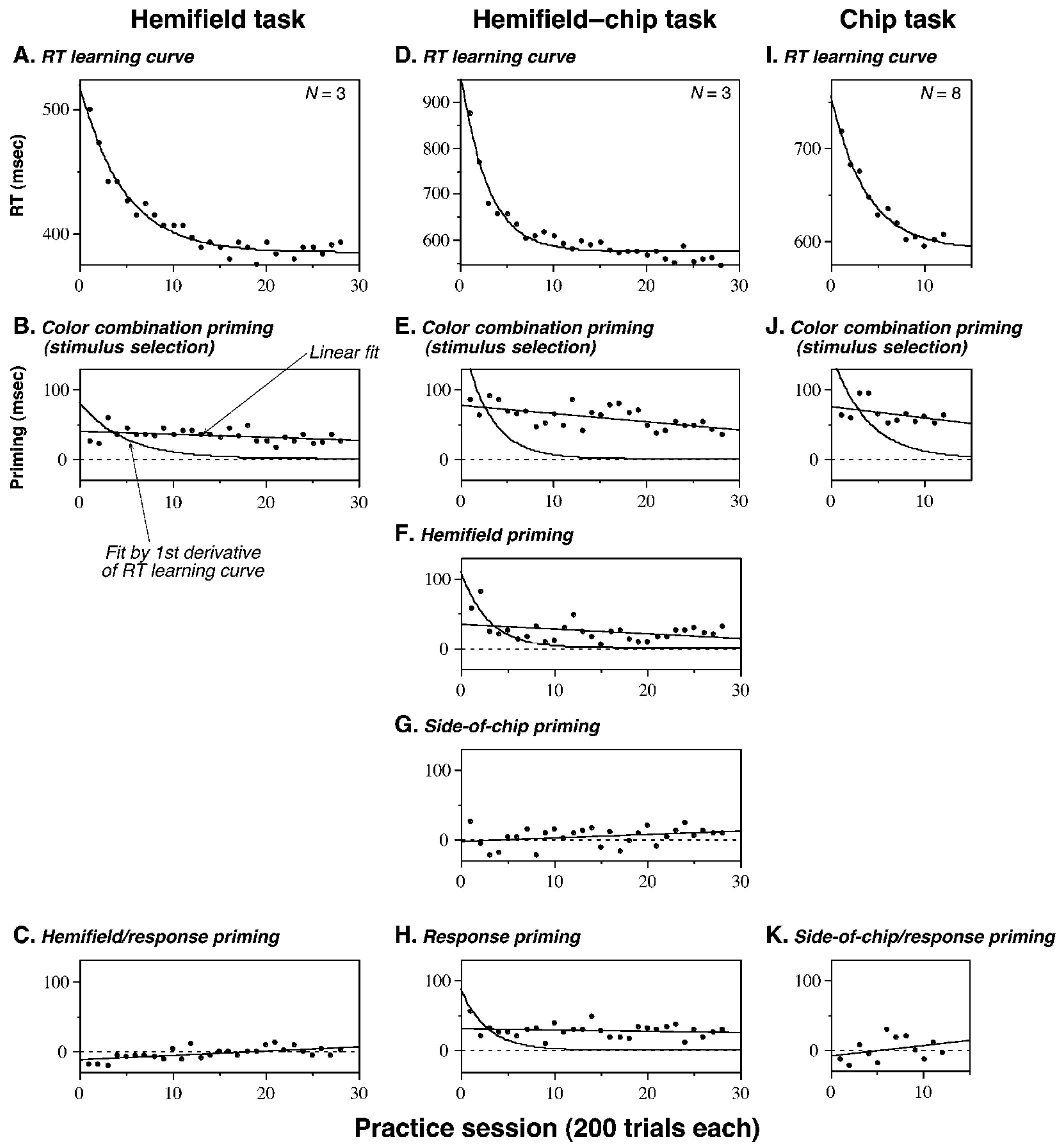

Figure 2. Mean response time (RT) and first-order sequential priming effects (in milliseconds) are shown as functions of practice sessions; the data are averaged across observers. A positive value of priming indicates RT facilitation—that is, the degree to which RTs were faster when a stimulus feature and/or a response was identical in the immediately preceding trial than when it was different. (A–C) The hemifield task: (A) mean RT, (B) color combination (stimulus selection) priming, and (C) hemifield/response priming. (D–H) The hemifield–chip task: (D) mean RT, (E) color combination (stimulus selection) priming, (F) hemifield priming, (G) side-of-chip priming, and (H) response priming. (I–K) The chip task: (I) mean RT, (J) color combination (stimulus selection) priming, and (K) side-of-chip/response priming. The RT learning curves (A, D, and I) are fit with exponential functions as an example. The priming data are fit with lines (linear regression) and with the first derivatives of the corresponding RT learning curves (the latter fits are not shown unless the priming effects exhibited decreasing trends over practice sessions). Note that for the hemifield task and the hemifield–chip task, the fits were analyzed separately for individual observers, using their entire practice sessions as shown in Tables 1–3, 5, and 6. Here, for simplicity, the observer-averaged data are shown only up to 28 sessions (for which data are available from all the observers).

the data were analyzed separately for individual observers. For the chip task, observer was used as the random effect. The color combination (stimulus selection) priming (range, 21–101 msec) was significant in all cases ($p < .05$, two-tailed, indicated by asterisks; Tables 2A, 3A and 4A). The hemifield priming (range, 17–40 msec) and the response priming (range, 12–41 msec) for the hemifield–chip task were significant in all cases (Tables 3B and 3D). Other priming effects were unreliable when averaged across practice sessions (Tables 2B, 3C, and 4B). The question of why certain priming effects were unreliable is addressed later.

Table 2
First-Order Priming Effects in the Hemifield Task

	Linear Fit		Fit With $C\cdot[dY_{\text{learn}}/dX]$ (RMSE Relative to Linear Fit)		Decay Rate Relative to Predicted		Overall Priming	
Observer	*r*	Slope (msec per Session)	Power	Exponential	Power Fit/$(d_0 + 1)$	Exponential Fit/α_0	Mean (msec)	*SE* (Session Based)
A. Color Combination (Stimulus Selection) Priming Over Sessions								
K.O.	−.372*	−0.740	2.816	2.776	0.025	0.028	40*	3.1
S.S.	−.644*	−0.521	2.219	2.006	0.221	0.175	31*	2.0
Y.S.	−.224	−0.205	2.505	2.411	0.050	0.032	21*	1.5
B. Hemifield Response Priming Over Sessions								
K.O.	.785*	1.917	n/a	n/a	n/a	n/a	0	3.8
S.S.	.212	0.078	n/a	n/a	n/a	n/a	−6	0.9
Y.S.	.142	0.132	n/a	n/a	n/a	n/a	0	1.5

*$p < .05$, two-tailed.

The critical question we asked was how priming changed over practice sessions. We first performed a linear regression analysis on each priming effect (priming vs. practice session). If the effects of practice and priming on RT were numerically additive, priming should remain constant over practice sessions. In other words, the correlation coefficient, *r*, should not be significantly different from zero. A significant negative correlation would indicate that priming decreased over practice sessions, whereas a significant positive correlation would indicate that priming increased over practice sessions. For example, Figure 2B shows that the color combination (stimulus selection) priming for the hemifield task remained relatively constant but decreased slightly over practice sessions. The corresponding analyses shown in Table 2A (under “Linear Fit”) indicate that the small decrease (at most, 0.8 msec per session; see under “Slope”) was significant for K.O. and S.S. (asterisks), but not for Y.S.

The continuous curve shown in Figure 2B is the fit with $C\cdot[dY_{\text{learn}}/dX]$ (*C* being the fitting parameter; least-squares fit, using Mathematica, Wolfram Research, Inc.)—that is, the fit with the first derivative of the learning curve shown in Figure 2A. Apparently, the fit is poor. To quantify how poorly $C\cdot[dY_{\text{learn}}/dX]$ fit practice-induced changes in priming, we computed the ratio of the root-mean square error (RMSE) for the $C\cdot[dY_{\text{learn}}/dX]$ fit to the RMSE for the linear fit. If the two fits were equally good (or bad), the ratio should be 1. If the $C\cdot[dY_{\text{learn}}/dX]$ fit was worse than the linear fit (as is evident in Figure 2B), the ratio should be greater than 1. As is shown under “Fit With $C\cdot[dY_{\text{learn}}/dX]$” in Table 2A, the fits with $C\cdot[dY_{\text{learn}}/dX]$ were 2–3 times as bad as the linear fits for

Table 3
First-Order Priming Effects in the Hemifield–Chip Task

	Linear Fit		Fit With $C\cdot[dY_{\text{learn}}/dX]$ (RMSE Relative to Linear Fit)		Decay Rate Relative to Predicted		Overall Priming	
Observer	*r*	Slope (msec per Session)	Power	Exponential	Power Fit/(d_0+1)	Exponential Fit/α_0	Mean (msec)	*SE* (Session Based)
A. Color Combination (Stimulus Selection) Priming Over Sessions								
K.O.	−.623*	−2.567	3.781	3.661	0.087	0.047	101*	6.4
S.S.	−.234*	−0.227	2.272	1.968	0.081	0.059	41*	2.4
Y.S.	−.424*	−0.680	2.033	1.907	0.187	0.099	33*	2.7
B. Hemifield Response Priming Over Sessions								
K.O.	−.411*	−1.875	1.223	1.065	0.321	0.193	40*	7.1
S.S.	−.143	−0.086	1.624	1.472	0.086	0.056	17*	1.5
Y.S.	.115	0.186	n/a	n/a	n/a	n/a	17*	2.7
C. Side-of-Chip Priming Over Sessions								
K.O.	.087	0.283	n/a	n/a	n/a	n/a	0	5.0
S.S.	.087	0.061	n/a	n/a	n/a	n/a	2	1.7
Y.S.	.056	0.094	n/a	n/a	n/a	n/a	12*	0.2
D. Response Priming Over Sessions								
K.O.	.248	0.536	n/a	n/a	n/a	n/a	41*	3.4
S.S.	−.534*	−0.440	1.401	1.222	0.307	0.282	15*	2.0
Y.S.	−.343	−0.671	1.241	1.268	0.277	0.142	12*	3.3

*$p < .05$, two-tailed.

Table 4
First-Order Priming Effects in the Chip Task

	Linear Fit		Fit With $C\cdot[dY_{\text{learn}}/dX]$ (RMSE Relative to Linear Fit)		Decay Rate Relative to Predicted		Overall Priming	
No. of Observers	r	Slope (msec per Session)	Power	Exponential	Power Fit/(d_0+1)	Exponential Fit/α_0	Mean (msec)	*SE* (Observer Based)
A. Color Combination (Stimulus Selection) Priming Over Sessions								
8	−.427	−1.721	3.647	3.219	0.087	0.099	64*	11.8
B. Side-of-Chip Priming Over Sessions								
8	.309	1.415	n/a	n/a	n/a	n/a	3	5.8

$^*p < .05$, two-tailed.

all 3 observers, whether the fits were based on power functions or exponential functions.

Note that $C\cdot[dY_{\text{learn}}/dX]$ fit the data poorly because the actual decay rate of the color combination (stimulus selection) priming was much slower than that predicted by $C\cdot[dY_{\text{learn}}/dX]$ (Figure 2B). To quantify this relationship, we refit the decay of priming over practice sessions, using $C\cdot[dY_{\text{learn}}/dX]$, but this time we also varied the decay rate as an additional fitting parameter; since $C\cdot[dY_{\text{learn}}/dX]$ takes the forms of $C\cdot(\text{X}+c_0)^{-D}$ (power function) and $C\cdot e^{-DX}$ (exponential function), C and D were both varied as the fitting parameters. The fitted values of D represent the actual decay rates of priming. Since the corresponding decay rates predicted by the RT learning curves are d_0+1 (power function) and α (exponential function), we computed the ratio of D to d_0+1 for power-function–based fits and the ratio of D to α for exponential-function–based fits. If the actual decay rate of priming was equal to that predicted by $C\cdot[dY_{\text{learn}}/dX]$, the ratio should be 1. If priming decayed more slowly than predicted (as is evident in Figure 2B), the ratio should be between 0 and 1. As is shown under "Decay Rate" in Table 2A, the color combination (stimulus selection) priming for the hemifield task decayed at rates only 3%–22% of those predicted by $C\cdot[dY_{\text{learn}}/dX]$ for the 3 observers.

The characteristics described above for the color combination (stimulus selection) priming for the hemifield task were obtained for all three tasks and for all priming effects (Tables 1–4; Figure 2). First, priming decayed rather slowly (or did not decay at all) over practice sessions. Linear decay of priming was significant in only 7 out of the 20 cases (see under "Linear Fit" in Tables 2–4), and even when significant, the decay was rather gradual (at most, 2.6 msec per session). Second, fits by $C\cdot[dY_{\text{learn}}/dX]$ were always worse than linear fits. The ratios of RMSE of $C\cdot[dY_{\text{learn}}/dX]$ fits to RMSE of linear fits were greater than 1 (1.1–3.8) in all cases (see under "Fit With $C\cdot[dY_{\text{learn}}/dX]$" in Tables 2–4). Third and most critical, fits by $C\cdot[dY_{\text{learn}}/dX]$ were poor because the actual priming decayed (if at all) much slower than predicted by $C\cdot[dY_{\text{learn}}/dX]$. The ratio of the actual rate of decay to the predicted rate of decay (based on dY_{learn}/dX) was less than 1 (0.03–0.32) in all cases[2] (see under "Decay Rate" in Tables 2–4). Note that fitting by $C\cdot[dY_{\text{learn}}/dX]$ cannot be meaningfully performed when priming showed an *increasing* trend over practice sessions (indicated as "n/a" in Tables 2–4). We also note that the results were consistent when the data from the 8 observers in the chip task were examined separately; the color combination priming (overall, significant for each observer) diminished significantly over sessions for only 1 observer, but with the decay rate being much less than that predicted from her learning curve.

In summary, whereas long-term practice substantially speeded RT in all the tasks (following learning curves that are well characterized by power and exponential functions), none of the first-order priming effects decayed at the rates predicted by the first derivatives of the corresponding learning curves. Priming either decayed much more gradually than predicted or did not decay at all.

Because stimulus characteristics and response were varied randomly from trial to trial, first-order priming effects (effects of immediate repetitions) were averaged across randomized prior repetition patterns. Thus, so long as individual trials contributed linearly, potential higher order sequential effects were averaged out when we computed first-order priming effects. However, higher order sequential effects are often nonlinear, especially when RSIs are relatively long, as in the present study (2–2.5 sec). This phenomenon is commonly known as expectancy (Kirby, 1976, 1980; Soetens, 1998; Soetens et al., 1985; Soetens, Deboeck, & Hueting, 1984).

Suppose that two stimuli (and two corresponding responses) are intermixed; we use R to indicate that the same stimulus/response was repeated in a pair of consecutive trials and A to indicate that the stimulus/response was changed (or alternated) in a pair of consecutive trials. Expectancy refers to the perceptual-cognitive system's tendency (under certain conditions) to become biased toward regular temporal patterns of repetitions and alternations. For example, following three consecutive repetitions in a row, RRR, another repetition, R, would be *expected*; following three alternations in a row, AAA, another alternation, A, would be *expected*. Thus, stimulus repetition, R, should result in faster RTs following RRR than following AAA; similarly, stimulus alternation, A, should result in faster RTs following AAA than following RRR. In general, if a particular prior repetition–alternation pattern biases expectancy more toward a repetition (reducing repetition RTs), it should bias expectancy more against an alternation (increasing alternation RTs), and vice versa. Thus, a cost–benefit type of negative correlation should occur between stimulus/response repetition RTs and stimulus/response alternation RTs for the current trial, depending on the preceding patterns of repetition–alternation (e.g., Audley, 1973; Kirby, 1972; Soetens, 1998; Soetens et al., 1985; Soetens, Deboeck, & Hueting, 1984). Mathematically, the strength of expectancy effects (the negative slope of the regression line) can vary orthogonally to the strength of first-order priming effects (the center of gravity of the correlation pattern being shifted toward faster repetition RT and slower alternation RT). However, expectancy and first-order priming might interact in the brain. If so, there is a remote possibility that first-order priming did decay as rapidly as predicted by the first derivative of the learning curves but that concurrent changes in expectancy almost exactly compensated for the decay.

To evaluate this possibility, we computed third-order expectancy effects—that is, we measured the effects of repetition–alternation patterns of the preceding 3 trials on the current repetition or alternation RT. This analysis was not performed for the chip task (our prior data) because the cued and prewarned trials had to be removed; too few trials were available to analyze higher order sequential effects. For observers K.O., S.S., and Y.S., we calculated expectancy effects (slopes of repetition–alternation functions) on correct trials (all 4 consecutive trials being correct) for each block of 200 trials and examined how they changed over practice sessions. If expectancy effects changed systematically, it is possible that they might have confounded the over-session changes in first-order priming effects. Alternatively, if expectancy effects did not change over practice sessions, we can rule out the possibility that changes in expectancy compensated for the actual reductions in first-order priming.

For each block of 200 trials, we pooled RTs for four pairs of repetition–alternation sequences: (RRR, RRA), (RAR, RAA), (ARR, ARA), and (AAR, AAA). Median RTs of these eight sequences (25 trials per sequence per block per observer, on average) were used to compute the correlation coefficient, r, and the slope of the regression line, ••A on ••R; slopes were computed using Brace's (1977) method of linear fit designed for cases in which both X and Y contain errors.

The overall repetition–alternation slopes were obtained by averaging across practice sessions. The standard error of the mean (*SE*) and statistical significance (asterisks indicating $p < .05$, two-tailed) were computed using practice session as the random effect (see under "Overall Expectancy" in Tables 5 and 6); the slopes from individual practice sessions were weighted by r^2 in computing the means and *SE*s so that more reliable slopes received higher weights. Note that most slopes were negative (15 out of 18 cases in Tables 5 and 6), indicating that expectancy trends were indeed present in most cases (not surprising, because we used relatively long RSIs). Though a prior study using a two-choice IRP found expectancy effects only for response-related repetitions (not for stimulus repetitions; Soetens, 1998), we found some significant expectancy effects for color combination (stimulus selection) priming, as well as for response-related priming. This difference may be due to the different paradigms being used to dissociate stimulus and response priming (IRP vs. our paradigm), to inclusion of a stimulus selection stage in our tasks, and/or to color pop-out being salient (e.g., Kirby, 1980) and highly task relevant in our study. Further research is needed to evaluate these and other possibilities.

Table 5
Expectancy Effects (Slopes of Repetition–Alternation Functions) in the Hemifield Task

		Overall Expectancy	
Observer	Linear Regression of Expectancy on Practice Session (R)	Mean	*SE* (Session Based)
A. Color Combination (Stimulus Selection) Priming Over Sessions			
K.O.	−.038	−0.687	0.443
S.S.	.024	−1.024*	0.099
Y.S.	.427*	−1.434*	0.184
B. Hemifield Response Priming Over Sessions			
K.O.	−.329	−0.647*	0.176
S.S.	−.201	0.072	0.145
Y.S.	−.060	−0.483	0.209

*$p < .05$, two-tailed.

The primary goal of this analysis was to determine how expectancy effects changed over practice sessions. We thus regressed the repetition–alternation slope on practice session (individual slopes were again weighted by their goodness of fit, r^2). Significant correlation would indicate that expectancy changed systematically over practice sessions. As is shown in Tables 5 and 6 (under "Linear Regression of Expectancy on Practice Session"), the linear correlation (labeled R) was significant in only 1 (color combination priming in the hemifield task for Y.S.) out of the 18 cases shown in Tables 5 and 6. To determine whether expectancy and first-order priming were directly related in this case (rather than being separately affected by practice), we examined their mutual correlation after variances explained by practice session were removed. The residual correlation ($r^2 = .071$) was not significant. We thus conclude that expectancy effects did not change over practice sessions, except for one case, but that even in that case, expectancy did not directly affect first-order priming.

Because expectancy effects (the major nonlinear contribution of higher order sequential effects) did not substantially change over practice sessions, we conclude that first-order priming effects indeed remained relatively constant over practice sessions (or decayed much more gradually than predicted by the first derivative of the learning curves).

Finally, we note that first-order response priming was negligible in the hemifield task and the chip task (Figures 2C and 2K; Tables 2B and 4B). This could be due

Table 6
Expectancy Effects (Slopes of Repetition–Alternation Functions) in the Hemifield–Chip Task

Observer	Linear Regression of Expectancy on Practice Session (R)	Overall Expectancy Mean	Overall Expectancy SE (Session Based)
A. Color Combination (Stimulus Selection) Priming Over Sessions			
K.O.	−.055	−0.148	0.817
S.S.	−.031	−0.996*	0.178
Y.S.	−.040	−0.668	0.595
B. Hemifield Response Priming Over Sessions			
K.O.	.095	0.213	0.357
S.S.	−.077	−0.310	0.178
Y.S.	.193	−0.586	0.313
C. Side-of-Chip Priming Over Sessions			
K.O.	−.190	−0.932*	0.442
S.S.	−.114	0.042	0.157
Y.S.	.099	−0.389	0.209
D. Response Priming Over Sessions			
K.O.	−.041	−0.184	0.221
S.S.	−.086	−0.431*	0.158
Y.S.	.008	−0.624*	0.170

*$p < .05$, two-tailed.

to first-order alternation effects. It is known that when RSI is relatively long and S–R compatibility is high in a two-choice task (as in the hemifield and chip tasks), people tend to anticipate the response to change from trial to trial (gambler's fallacy; see e.g., Jarvik, 1951; Wagenaar, 1972), reducing or even reversing first-order repetition effects (e.g., Entus & Bindra, 1970; Kirby, 1976; Soetens, 1998; Soetens et al., 1985; Soetens, Deboeck, & Hueting, 1984; see also Goolsby & Suzuki, 2001). Our results are consistent because when the S–R translation was made complex in the hemifield–chip task, robust first-order response priming was obtained (Figure 2H; Table 3D). The fact that the hemifield/response priming and the side-of-chip/response priming had slightly increasing trends over practice sessions (Figures 2C and 2K; Tables 2B and 4B) might be due to the fact that first-order alternation effects tend to gradually diminish with practice (e.g., Soetens et al., 1985).

DISCUSSION

We investigated whether or not long-term practice and sequential priming modulated a common set of processing stages through the same mechanism. We examined how the magnitudes of various priming effects changed over practice sessions while overall RT diminished as exponential or power functions. All first-order priming effects remained relatively constant over practice sessions or diminished (approximately linearly) at much slower rates than would be predicted from the shape (first derivative) of the corresponding RT learning curves. Expectancy effects also remained relatively constant over practice sessions, ruling out the possibility that practice-induced changes in expectancy somehow compensated for rapid decays of first-order priming through nonlinear interactions.

As was discussed in the introduction, the fact that all priming effects diminished substantially more slowly than predicted by dY_{learn}/dX (if they did at all) suggests two possible interpretations. First, practice primarily affected a processing stage untapped by any of the priming effects measured. This interpretation is unlikely because the three tasks used allowed us to monitor priming effects on the critical processing stages of (1) stimulus selection (color combination priming in all tasks), (2) stimulus identification/categorization (hemifield and chip priming in the hemifield–chip task), and (3) response selection (hemifield/response priming in the hemifield task, chip/response priming in the chip task, and response priming in the hemifield–chip task). As was discussed earlier, all these processing stages are known to be influenced by both long-term practice and sequential priming. It is thus likely that both practice and priming facilitated largely overlapping sets of processing stages, but through separate mechanisms; mechanisms underlying practice effects scaled inputs by dY_{learn}/dX, but mechanisms underlying priming effects did not.

One assumption we made in deriving these conclusions was that practice effects on component processing stages make *linear contributions* to observed RT learning curves. If the decay of each priming effect was scaled by the first derivative of the corresponding learning curve, the obtained slow decays of individual priming effects would imply slow learning rates for the corresponding component learning curves. As is shown in the Appendix, linear contributions from these slow component learning curves could not generate an overall RT learning curve whose learning rate is greater than those of the individual component learning curves. Contrary to this prediction, the observed learning curves had much greater learning rates than did those implied by any of the priming decay functions (note that the learning rates referred to here are those beyond constant scaling factors, represented by ds in power functions and αs in exponential functions). We thus concluded, by reasoning by contradiction, that practice and priming modulated RT through separate mechanisms.

However, what if the component learning curves contributed nonlinearly? For example, multiplicative contributions from slow power/exponential learning curves could generate an overall learning curve whose learning rate is greater than those for the component learning curves because ds and αs would add when power/exponential functions are multiplied. It is thus possible that practice and priming operated through the same slow-learning mechanism but that the component learning curves contributed multiplicatively to generate the observed steep learning curves. It is, however, difficult to postulate a simple mechanism in which RTs for the component processing stages contribute multiplicatively to the overall RT. For example, if the component processing stages were executed sequentially (e.g., stimulus selection → stimulus identification/categorization → response selection → response), their contributions to the overall RT should be additive (e.g., Newell & Rosenbloom, 1981; Sternberg, 1969). If any of the component stages were executed in parallel, the contributions should then be subadditive. In fact, most learning models that postulate contributions from component processing stages assume the contributions to be linear, whereas none assumes the contributions to be multiplicative (e.g., Kirsner & Speelman, 1996; Newell & Rosenbloom, 1981; Rickard, 1997; see Heathcote et al., 2000, for a review). Thus, although we cannot rule out the possibility of multiplicative contributions, we tentatively conclude that this is an unlikely alternative.

Relation to Other Studies

Soetens, Deboeck, and Hueting (1984) examined automatic sequential priming effects on a choice RT task using short RSIs and high S–R compatibility; observers pressed the left key (using the left hand) when "0" was presented to the left of a central vertical line and pressed the right key (using the right hand) when "0" was presented to the right of a central vertical line. In their Experiment 3, sequential priming effects (due to repetitions/alternations of stimulus location/response) were monitored as observers practiced the task over 7,000 trials (1,000 trials per session). The average RT decreased clearly following a power/exponential function, whereas the first-order priming diminished linearly or in a slightly accelerating manner (we obtained the learning curve and the first-order priming from Figure 5 of Soetens, Deboeck, and Hueting, 1984), demonstrating that first-order priming was not constrained by the shape of the RT learning curve. Hence, we extended Soetens, Deboeck, and Hueting's results, using three tasks that differed in the level of difficulty (RTs for the initial sessions ranging from 450 to 1,100 msec), using long RSIs (range, 2–2.5 sec), using simple and complex S–R mappings (in the hemifield and chip tasks and in the hemifield–chip task, respectively), and using priming that presumably affected different processing stages (e.g., stimulus selection, stimulus property identification, and response selection).

The present results also complement the previous studies that compared effects of long-term practice and first-order sequential priming, using a transfer paradigm. In a transfer paradigm, potential relationships between practice and priming are assessed by examining whether effects of practice and priming transfer to a similar set of new conditions (e.g., new stimuli, new S–R mapping, and new mode of response).

To examine transfer of practice effects, participants are typically trained extensively in one condition (A) and later tested in a new condition (B). A complete transfer of practice (i.e., training in Condition A improves performance in a new Condition B as much as it improves performance in Condition A) suggests that practice optimizes a common processing stage that encompasses both Conditions A and B. A poor transfer of practice (i.e., training in Condition A does not improve performance in Condition B) suggests that practice optimizes a processing stage involved in Condition A but not involved in Condition B. A partial transfer (i.e., training in Condition A only partially improves performance in Condition B) suggests that practice optimizes a processing stage primarily involved in Condition A and only partially involved in Condition B. Thus, by choosing Conditions A and B appropriately, properties of the processing stages affected by long-term practice in Condition A can be inferred.

Similarly, processing stages affected by first-order sequential priming can be assessed by considering trial $n-1$ as a *practice* phase and trial n as a *transfer* phase. A complete transfer of priming corresponds to the case in which a prior Condition A trial benefits a current Condition B trial as much as it benefits a current Condition A trial. The benefits are measured relative to the cases in which either an A trial or a B trial is preceded by an apparently unrelated C trial (which corresponds to an arbitrary pretraining experience in long-term practice). A poor transfer of priming (from Condition A to Condition B) corresponds to the case in which a prior A trial does not benefit a current B trial (RT for a B trial being the same whether it is preceded by an A trial or by a C trial). A partial transfer corresponds to cases in which a prior A trial benefits a current B trial, but not as much as it benefits a current A trial.

Thus, processing stages affected by long-term practice and sequential priming can be compared by comparing their transfer characteristics under various conditions. This strategy was adopted by Pashler and Baylis (1991a, 1991b) and Campbell and Proctor (1993). Familiar symbols (digits, letters, and other keyboard symbols) were assigned to three fingers; the task was simply to flex the correct finger (to press the correct key) upon stimulus presentation. Transfers of long-term practice and sequential priming were examined under various conditions that differed in the following aspects: (1) whether same stimulus categories were consistently mapped to same responses or to different responses and (2) whether category-to-response mapping, response mapping (e.g., mapping to spatial positions and/or to fingers), mode of response (e.g., vocal vs. manual), and/or surface features of the symbols (e.g., case or color of letters) were preserved or altered at transfer. The overall results suggested that both long-term practice and sequential priming strengthened *high-level* connections between a category representation (e.g., digits) and an abstract response representation (e.g., left, middle, and right, rather than specific fingers) but also strengthened *low-level* connections between symbol representations and less abstract response representations (e.g., weakly hand specific). Although long-term practice appeared to strengthen the category-to-response connections relatively more strongly, whereas sequential priming appeared to strengthen the stimulus-to-response connections relatively more strongly, the data were inconclusive.

The transfer results thus suggested that practice and priming primarily affected a common set of behaviorally defined processing stages. In contrast, our present results suggest that practice (scaled by dY_{learn}/dX) and priming (unconstrained by dY_{learn}/dX) operate through different mechanisms even if they affect a common set of processing stages. These complementary findings together suggest that each processing stage may possess a separate mechanism for long-term and short-term plasticity.

Finally, if practice-induced decline in sequential priming is not mediated by whatever mechanisms underlie the exponential/power functions of the RT learning curve (see Heathcote et al., 2000, for a review), what might be causing the gradual reduction of sequential priming over practice sessions? In the case of color combination priming, its slow decay might be due to gradual development of a general mechanism for odd-color detection. Initially, a green target might be detected among red distractors by using a mechanism that detects green against red, whereas a red target might be detected among green distractors using a mechanism that detects red against green. The color combination priming could thus be due to short-term potentiation of specific large-scale color-opponent mechanisms. These mechanisms might be mediated by color-tuned cells in V4 with large suppressive surrounds (up to 30° of visual angle), which are responsive to color contrasts (e.g., Desimone & Schein, 1987; Desimone, Schein, Moran, & Ungerleider, 1985). However, with extensive practice on the task in which red and green targets appear randomly across trials, a more general representation might develop. In this emergent representation, an odd-colored item might be highlighted regardless of whether it is red or green. As the process of target selection becomes more reliant on this general odd-color representation, rather than on color-specific (red-against-green and green-against-red) representations, the color combination priming effects might diminish. Gradual decays of other priming effects we examined, however, cannot be explained as readily by postulating emergent representations. Nevertheless, it might be that gradual reductions in sequential priming in general require development of new representations or processing strategies; otherwise, priming effects might remain unchanged. It would be interesting to investigate this speculation in future research.

To conclude, the present results suggest that the effectiveness of sequential priming (decaying slowly and approximately linearly over practice sessions) is not constrained by changes in the effectiveness of practice over training sessions (exponential/power form of RT learning curves). This dissociation was found for several priming effects presumably influencing different behaviorally defined processing stages (stimulus selection, stimulus property identification, and response selection). A plausible interpretation based on the present results and the existing literature is that short-term sequential priming and long-term practice modulate a largely common set of processing stages, but through relatively separate and noninteractive mechanisms. It might be that various processing stages involved in speeded S–R tasks in general possess mechanisms for both short-term and long-term plasticity so that people can take advantage of both short-term and long-term regularity in their behavioral environment.

REFERENCES

Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences*, **90**, 5718-5722.

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

Ahissar, M., Laiwand, R., Kozminsky, G., & Hochstein, S. (1998). Learning pop-out detection: Building representations for conflicting target–distractor relationships. *Vision Research*, **38**, 3095-3107.

Audley, R. J. (1973). Some observations on theories of choice reaction time: Tutorial review. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 509-545). New York: Academic Press.

Ball, K., & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science*, **218**, 697-698.

Bertelson, P. (1965). Serial choice reaction-time as a function of response versus signal-and-response repetition. *Nature*, **206**, 217-218.

Brace, R. A. (1977). Fitting straight lines to experimental data. *American Journal of Physiology*, **233**, R94-R99.

Campbell, K. C., & Proctor, R. W. (1993). Repetition effects with categorizable stimulus and response sets. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 1345-1362.

Czerwinski, M., Lightfoot, N., & Shiffrin, R. M. (1992). Automatization and training in visual search. *American Journal of Psychology*, **105**, 271-315.

Desimone, R., & Schein, S. J. (1987). Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *Journal of Neurophysiology*, **57**, 835-868.

Desimone, R., Schein, S. J., Moran, J., & Ungerleider, L. G. (1985). Contour, color, and shape analysis beyond the striate cortex. *Vision Research*, **25**, 441-452.

Entus, A., & Bindra, D. (1970). Common features of the "repetition" and "same–different" effects in reaction time experiments. *Perception & Psychophysics*, **7**, 143-148.

Fahle, M. (1994). Human pattern recognition: Parallel processing and perceptual learning. *Perception*, **23**, 411-427.

Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature*, **287**, 43-44.

Goolsby, B. A., & Suzuki, S. (2001). Understanding priming of color-singleton search: Roles of attention at encoding and "retrieval." *Perception & Psychophysics*, **63**, 929-944.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review*, **7**, 185-207.

Hillstrom, A. P. (2000). Repetition effects in visual search. *Perception & Psychophysics*, **62**, 800-817.

Hillstrom, A. P., & Logan, G. D. (1998). Decomposing visual search: Evidence of multiple item-specific skills. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1385-1398.

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology*, **41**, 291-297.

Kirby, N. H. (1972). Sequential effects in serial reaction time. *Journal of Experimental Psychology*, **96**, 32-36.

Kirby, N. H. (1976). Sequential effects in two-choice reaction time: Automatic facilitation or subjective expectancy? *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 567-577.

Kirby, N. H. (1980). Sequential effects in choice reaction time. In A. T. Welford (Ed.), *Reaction times* (pp. 129-172). London: Academic Press.

Kirsner, K., & Speelman, C. (1996). Skill acquisition and repetition priming: One principle, many processes? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 563-575.

Kornblum, S. (1973). Sequential effects in choice reaction time: A tutorial review. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 259-288). New York: Academic Press.

Kröse, B. J. A., & Julesz, B. (1989). The control and speed of shifts of attention. *Vision Research*, **29**, 1607-1619.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, **22**, 1-35.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition*, **22**, 657-672.

Maljkovic, V., & Nakayama, K. (2000). Priming of pop-out: III. A short-term implicit memory system beneficial for rapid target selection. *Visual Cognition*, **7**, 571-595.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the power law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

Pashler, H., & Baylis, G. (1991a). Procedural learning: 1. Locus of practice effects in speeded choice tasks. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 20-32.

Pashler, H., & Baylis, G. (1991b). Procedural learning: 2. Intertrial repetition effects in speeded-choice tasks. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 33-48.

Polat, U., & Sagi, D. (1994). Spatial interactions in human vision: From near to far via experience-dependent cascades of connections. *Proceedings of the National Academy of Sciences*, **91**, 1206-1209.

Rabbitt, P. M. (1968). Repetition effects and signal classification strategies in serial choice-response tasks. *Quarterly Journal of Experimental Psychology*, **20**, 232-240.

Rabbitt, P., Cumming, G., & Vyas, S. (1979). Modulation of selective attention by sequential effects in visual search tasks. *Quarterly Journal of Experimental Psychology*, **31**, 305-317.

Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, **126**, 288-311.

Rovamo, J., & Virsu, V. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research*, **37**, 475-494.

Shiffrin, R. M., & Lightfoot, N. (1997). Perceptual learning of alphanumeric-like characters. In R. L. Goldstone, D. L. Medin, & P. G. Schyns (Eds.), *Perceptual learning* (pp. 45-81). San Diego: Academic Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, **84**, 127-190.

Simon, J. R., & Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus–response correspondence and age. *Ergonomics*, **6**, 99-105.

Sireteanu, R., & Rettenbach, R. (2000). Perceptual learning in visual search generalizes over tasks, locations, and eyes. *Vision Research*, **40**, 2925-2949.

Smith, M. C. (1968). Repetition effect and short-term memory. *Journal of Experimental Psychology*, **77**, 435-439.

Soetens, E. (1998). Localizing sequential effects in serial choice reaction time with the information reduction procedure. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 547-568.

Soetens, E., Boer, L. C., & Hueting, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 598-616.

Soetens, E., Deboeck, M., & Hueting, J. E. (1984). Automatic aftereffects in two-choice reaction time: A mathematical representation of some concepts. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 581-598.

Soetens, E., Deboeck, M., Hueting, J. [E.], & Merckx, H. (1984). Spatial compatibility with a two-dimensional stimulus arrangement. *Bulletin of the Psychonomic Society*, **22**, 125-128.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II* (pp. 276-315). Amsterdam: North-Holland.

Treisman, A., Vieira, A., & Hays, A. (1992). Automaticity and preattentive processing. *American Journal of Psychology*, **105**, 341-362.

Vervaeck, K. R., & Boer, L. C. (1980). Sequential effects in two-choice reaction time: Subjective expectancy and automatic aftereffect at short response–stimulus intervals. *Acta Psychologica*, **44**, 175-190.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin*, **77**, 65-72.

Welford, A. T. (1980). *Reaction times*. London: Academic Press.

NOTES

1. As Maljkovic and Nakayama (1994) pointed out, the side-of-chip discrimination they used (which is adopted here) requires a relatively high acuity judgment. Previous work has shown that acuity diminishes, as a function of visual eccentricity, faster along the vertical axis than along the horizontal axis (Kröse & Julesz, 1989; Rovamo & Virsu, 1979). We thus adopted the horizontally elongated ellipse used by Maljkovic and Nakayama in order to approximately equalize the acuity at all possible target positions.

2. We note that even when the curvature of the power and exponential functions, D, was allowed to vary as the fitting parameter, they still did not produce substantially better fits than did linear functions. Overall, the average RMSE was only 0.45% ($SD = 6.8\%$) less for the power and exponential fits than for the linear fits; in fact, the linear fits were better than the power and exponential fits in 12 out of the 22 cases (in which priming had decreasing trends over sessions—i.e., negative rs under "Linear Fit" in Tables 2–4). Thus, the gradual decay of priming, when it occurred, was primarily linear.

APPENDIX

We will show that if the overall (observed) RT learning curve consists of linear contributions from the component learning curves (e.g., representing practice effects on individual processing stages), the observed learning curve should reduce its RT faster than the slowest component learning curve and slower than the fastest component learning curve (when learning is measured as proportional reductions in RT).

We assume that the observed RT learning curve, $Y(X)$, is well approximated by a linear combination (weighted sum) of the component RT learning curves, $y_i(X)$—that is,

$$Y(X) \approx \sum q_i\{y_i(X)\}, \tag{A1}$$

where q_i are the weighting factors. For power functions, Equation A1 becomes

$$Y(X) = a_0 + b_0(X + c_0)^{-d_0} \approx \sum q_i\left\{a_i + b_i(X + c_i)^{-d_i}\right\}, \tag{A2}$$

where subscript 0 indicates parameters for the observed RT learning curve and subscripts i indicates parameters for the component learning curves. Similarly, for exponential functions, we have

$$Y(X) = f_0 + g_0 \cdot e^{-\alpha_0 X} \approx \sum q_i\left\{f_i + g_i \cdot e^{-\alpha_i X}\right\}. \tag{A3}$$

The learning rate in terms of proportions of reduction in RT, *proportional learning rate* (*PLR*) for the observed RT learning curve, PLR_0, is given by

$$PLR_0 = \frac{1}{Y(X) - Asymp.} \cdot \frac{d\{Y(X) - Asymp.\}}{dX}. \tag{A4}$$

For power functions, Equation A4 becomes (upon substitution of Equation A2 into Equation A4)

$$PLR_0 = \frac{-b_0 d_0 (X + c_0)^{-(d_0+1)}}{b_0 (X + c_0)^{-d_0}} - \frac{d_0}{X + c_0} \tag{A5}$$

(in terms of parameters of the observed learning curve), or

$$PLR_0 \approx \frac{\sum q_i\left\{-b_i d_i (X + c_i)^{-(d_i+1)}\right\}}{\sum q_i\left\{b_i (X + c_i)^{-d_i}\right\}} \tag{A6}$$

(in terms of parameters of the component learning curves). The latter expression can be simplified if we assume that prepractice levels of training for the different component processing stages are equivalent—that is, $c_i = c_0$ (as will be shown later, this assumption is not necessary for exponential functions). Upon substitution of $c_i = c_0$ into Equation A6, we get

$$PLR_0 \approx -\frac{1}{X + c_0} \cdot \frac{\sum q_i b_i d_i (X + c_0)^{-d_i}}{\sum q_i b_i (X + c_0)^{-d_i}}. \tag{A7}$$

If the slowest component learning curve had $d_{\min}$ and the fastest component learning curve had $d_{\max}$, we have (by subtracting $-[d_{\min}/(X + c_0)]$ or $-[d_{\max}/(X + c_0)]$ from Equation A7, where $0 \leq d_{\min} \leq d_i \leq d_{\max}$)

$$\frac{d_{\min}}{X + c_0} \leq \left|PLR_0\right| \leq \frac{d_{\max}}{X + c_0}, \tag{A8}$$

indicating that the proportional learning rate of the observed learning curve is bound by those of the slowest and the fastest component learning curves. Equation A8 simplifies to $d_{\min} \leq d_0 \leq d_{\max}$ (upon substitution of Equation A5).

Similarly for exponential functions, the observed proportional learning rate is given by (upon substitution of Equation A3 into Equation A4)

$$PLR_0 = \frac{-g_0 \alpha_0 \cdot e^{-\alpha_0 X}}{g_0 \cdot e^{-\alpha_0 X}} = -\alpha_0 \tag{A9}$$

(in terms of parameters of the observed learning curve), or

$$PLR_0 \approx \frac{\sum q_i\left\{-g_i \alpha_i \cdot e^{-\alpha_i X}\right\}}{\sum q_i\left\{g_i \cdot e^{-\alpha_i X}\right\}} \tag{A10}$$

(in terms of parameters of the component learning curves). If the slowest component learning curve had $\alpha_{\min}$ and the fastest component learning curve had $\alpha_{\max}$, we have (by subtracting $-\alpha_{\min}$ or $-\alpha_{\max}$ from Equation A10, where $0 \leq \alpha_{\min} \leq \alpha_i \leq \alpha_{\max}$)

$$\alpha_{\min} \leq \left|PLR_0\right| \leq \alpha_{\max}, \tag{A11}$$

APPENDIX (Continued)

again, indicating that the proportional learning rate of the observed learning curve is bound by those of the slowest and the fastest component learning curves. Equation A11 simplifies to $\alpha_{\min} \le \alpha_0 \le \alpha_{\max}$ (upon substitution of Equation A9).

Thus, if the observed RT learning curve is a linear combination of the component learning curves (representing practice effects on the component processing stages), the *PLR* of the observed learning curve, represented by d_0 (for power functions) and α_0 (for exponential functions), is bound by the *PLR* of the slowest and the fastest component learning curves—that is, $d_{\min} \le d_0 \le d_{\max}$ (for power functions) and $\alpha_{\min} \le \alpha_0 \le \alpha_{\max}$ (for exponential functions).

We note that a sum of component power functions yields an overall power function only when $c_i \approx c_0$ and $d_i \approx d_0$ (see Equation A2), and a sum of component exponential functions yields an overall exponential function only when $\alpha_i \approx \alpha_0$ (see Equation A3; or else, the weighting factors, q_i, must be appropriately chosen). We know from our data that all of our overall RT learning curves were well fit by power and exponential functions (Table 1). Furthermore, since the obtained over-session priming decay functions are no better fit by linear and other functions than by power and exponential functions (with d and α varied as fitting parameters), the implied component learning curves also appear to be appropriately described by power and exponential functions for our purposes (with small d and α). Finally, we emphasize that the overall proportional learning rate, PLR_0, for a linear combination of component learning curves (that are monotonically decaying functions) is always bound by the slowest ($plr_{\min}$) and the fastest ($plr_{\max}$) component learning curves regardless of their exact mathematical form.

In general, the proportional learning rate for the overall learning curve is given by

$$PLR_0 = \frac{1}{Y(X)-A}\cdot\frac{d\{Y(X)-A\}}{dX} = \frac{1}{\sum\{y_i(X)-a_i\}}\cdot\sum\frac{d\{y_i(X)-a_i\}}{dX},$$

where $Y(X) = \sum y_i(X)$ and A and a_i are the corresponding asymptotes. We let *PLR* for the fastest component learning curve be

$$plr_{\max} = \frac{1}{y_{\max}(X)-a_{\max}}\cdot\frac{d\{y_{\max}(X)-a_{\max}\}}{dX} \ge \frac{1}{y_i(X)-a_i}\cdot\frac{d\{y_i(X)-a_i\}}{dX} \qquad \text{(A12)}$$

and *PLR* for the slowest component learning curve be

$$plr_{\min} = \frac{1}{y_{\min}(X)-a_{\min}}\cdot\frac{d\{y_{\min}(X)-a_{\min}\}}{dX} \le \frac{1}{y_i(X)-a_i}\cdot\frac{d\{y_i(X)-a_i\}}{dX}. \qquad \text{(A13)}$$

Then, we have

$$PLR_0 - plr_{\max} = \frac{1}{\sum\{y_i(X)-a_i\}}\cdot\sum\frac{d\{y_i(X)-a_i\}}{dX} - \frac{1}{y_{\max}(X)-a_{\max}}\cdot\frac{d\{y_{\max}(X)-a_{\max}\}}{dX}. \qquad \text{(A14)}$$

The right side of Equation A14 can be reduced to

$$\frac{1}{\sum\{y_i(X)-a_i\}}\left[\sum\frac{d\{y_i(X)-a_i\}}{dX} - \sum\{y_i(X)-a_i\}\frac{1}{y_{\max}(X)-a_{\max}}\frac{d\{y_{\max}(X)-a_{\max}\}}{dX}\right],$$

which equals

$$\frac{1}{\sum\{y_i(X)-a_i\}}\left[\sum\{y_i(X)-a_i\}\left\{\frac{1}{y_i(X)-a_i}\frac{d\{y_i(X)-a_i\}}{dX}\right\} - \sum\{y_i(X)-a_i\}\left\{\frac{1}{y_{\max}(X)-a_{\max}}\frac{d\{y_{\max}(X)-a_{\max}\}}{dX}\right\}\right] \le 0$$

(due to Equation A12). Thus, $PLR_0 \le plr_{\max}$. Similarly, we have

$$PLR_0 - plr_{\min} = \frac{1}{\sum\{y_i(X)-a_i\}}\cdot\sum\frac{d\{y_i(X)-a_i\}}{dX} - \frac{1}{y_{\min}(X)-a_{\min}}\cdot\frac{d\{y_{\min}(X)-a_{\min}\}}{dX}. \qquad \text{(A15)}$$

The right side of Equation A15 can be reduced to

$$\frac{1}{\sum\{y_i(X)-a_i\}}\left[\sum\{y_i(X)-a_i\}\left\{\frac{1}{y_i(X)-a_i}\frac{d\{y_i(X)-a_i\}}{dX}\right\} - \sum\{y_i(X)-a_i\}\left\{\frac{1}{y_{\min}(X)-a_{\min}}\frac{d\{y_{\min}(X)-a_{\min}\}}{dX}\right\}\right] \ge 0$$

(due to Equation A13). Thus, $PLR_0 \ge plr_{\min}$. We thus have $plr_{\min} \le PLR_0 \le plr_{\max}$.

(Manuscript received October 30, 2000;
revision accepted for publication October 10, 2002.)