Psychonomic Bulletin & Review
2009, 16 (6), 1026-1036
doi:10.3758/16.6.1026

Brief Reports

A diffusion model decomposition of the practice effect

Gilles Dutilh
University of Amsterdam, Amsterdam, The Netherlands

Joachim Vandekerckhove and Francis Tuerlinckx
Leuven University, Leuven, The Netherlands

and

Eric-Jan Wagenmakers
University of Amsterdam, Amsterdam, The Netherlands

When people repeatedly perform the same cognitive task, their mean response times (RTs) invariably decrease. The mathematical function that best describes this decrease has been the subject of intense debate. Here, we seek a deeper understanding of the practice effect by simultaneously taking into account the changes in accuracy and in RT distributions with practice, both for correct and error responses. To this end, we used the Ratcliff diffusion model, a successful model of two-choice RTs that decomposes the effect of practice into its constituent psychological processes. Analyses of data from a 10,000-trial lexical decision task demonstrate that practice not only affects the speed of information processing, but also response caution, response bias, and peripheral processing time. We conclude that the practice effect consists of multiple subcomponents, and that it may be hazardous to abstract the interactive combination of these subcomponents in terms of a single output measure such as mean RT for correct responses. Supplemental materials may be downloaded from http://pbr.psychonomic-journals.org/content/supplemental.

When people repeatedly perform the same task, their performance becomes fast, accurate, and relatively effortless. For example, you are able to read this text quickly, virtually without errors, and, hopefully, without investing too much effort. The difference between your performance now and when you first learned to read is staggering; from a slow, error-prone, and effortful endeavor, your reading has matured into automatized skill.

Traditionally, researchers in the field of skill acquisition have quantified the effect of practice primarily in terms of a reduction in the time to execute a given task (i.e., response time or RT; Logan, 1992; Newell & Rosenbloom, 1981; Woodworth & Schlosberg, 1954). Almost every study has shown that the RT benefits due to practice are greatest at the start of training and then slowly diminish over time.

This ubiquitous result, many researchers have argued, is best captured by a *power function* that relates the mean RT to practice via the equation

$$\mathrm{MRT} = a + bN^{-c}, \tag{1}$$

where MRT is the mean RT for correct responses, a quantifies asymptotic performance, b quantifies the difference between initial and asymptotic performance, N represents the amount of practice, and c is the rate parameter that determines the shape of the power law. Empirical support for the power function relation between RT and practice has been reported across a range of tasks—for instance, in cigar rolling and maze solving (Crossman, 1959), fact retrieval (Pirolli & Anderson, 1985), and a variety of standard psychological tasks (Logan, 1992). Support for the power function relation has appeared so strong that the relation is often referred to as a law (e.g., "the ubiquitous law of practice," Newell & Rosenbloom, 1981, p. 3).

Nevertheless, some researchers have questioned whether the speedup with practice is really governed by a power function. In particular, Heathcote, Brown, and Mewhort (2000) argued that the power law is an artifact of averaging practice functions over participants (see also R. B. Anderson & Tweney, 1997; Myung, Kim, & Pitt, 2000). Heathcote et al. showed that for the data of many experiments on skill acquisition, individual learning

G. Dutilh, gilles.dutilh@gmail.com

© 2009 The Psychonomic Society, Inc.

curves were better described by an *exponential function* that relates mean RT to practice via the equation

$$\mathrm{MRT} = a + b\exp(-cN), \qquad (2)$$

where the interpretation of the parameters is the same as in Equation 1.

Regardless of the specific shape of the function that relates the amount of practice to mean RT, the previous discussion illustrates that most empirical studies on the practice effect have focused on the decrease in mean RTs for correct responses. By doing so, the field has largely neglected two other important sources of information—namely, accuracy (i.e., proportion of correct responses) and the distribution of RTs (e.g., spread and skewness). Those researchers who have taken response accuracy into account have tended to ignore RTs (e.g., Dosher & Lu, 2007; but see Nosofsky & Alfonso-Reese, 1999) or to present both RT and accuracy as separate output variables—even though it is well known that RT and accuracy are intimately related (see, e.g., Forstmann et al., 2008; Schouten & Bekker, 1967).

In this article, we seek a detailed understanding of the effect of practice by taking into account simultaneously the changes in accuracy and in RT distributions, for both correct and error responses. To do so, we follow Ratcliff, Thapar, and McKoon (2006) in applying the Ratcliff diffusion model (e.g., Ratcliff, 1978; Ratcliff & McKoon, 2008; Wagenmakers, 2009) to the field of automatization in cognitive tasks. The application of the diffusion model allows us to use all of the information in the data and to decompose the practice effect into its constituent psychological processes.

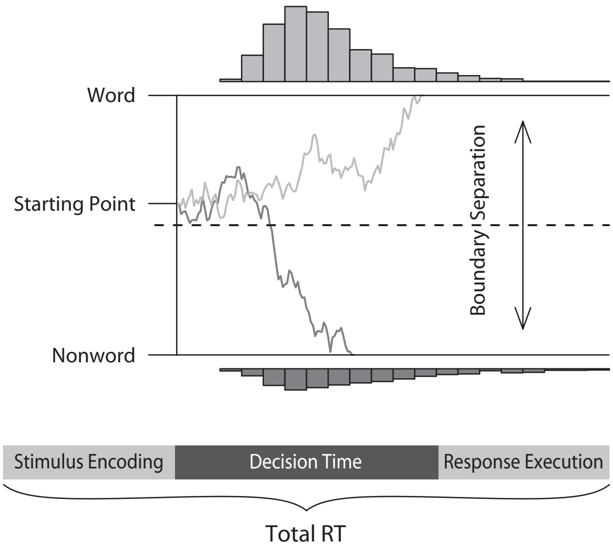

Figure 1. The diffusion model as it applies to the lexical decision task. A word stimulus (not shown) is presented, and two example paths represent the accumulation of evidence resulting in one correct response (light line) and one error response (dark line). Repeated application of the diffusion process yields histograms of both correct responses (upper histogram) and incorrect responses (lower histogram). As is evident from the histograms, the correct (upper) word boundary is reached more often than the incorrect (lower) nonword boundary. The total RT consists of the sum of a decision component, modeled by the noisy accumulation of evidence, and a nondecision component that represents the time needed for processes such as stimulus encoding and response execution.

THE RATCLIFF DIFFUSION MODEL

Here we describe the Ratcliff diffusion model as it applies to the lexical decision task, in which participants have to decide quickly whether a letter string is a word, such as *party*, or a nonword, such as *drapa* (Wagenmakers, Ratcliff, Gomez, & McKoon, 2008).

The core of the model is the Wiener diffusion process, which describes how the relative evidence for two response alternatives accumulates over time. The meandering light gray line in Figure 1 illustrates the continuous and noisy accumulation of evidence for a word response over a nonword response when a word is presented. When the amount of diagnostic evidence for one of the response options reaches a predetermined response threshold (i.e., one of the horizontal boundaries in Figure 1), the corresponding response is initiated. The dark gray line shows how the noise inherent in the accumulation process can sometimes cause the process to end up at the wrong (here, nonword) response boundary.

The diffusion model provides a detailed and comprehensive account of performance in speeded two-choice tasks (Ratcliff, 1978; Ratcliff & McKoon, 2008; Wagenmakers, 2009). In the model, unobserved psychological processes—represented by parameters—give rise to the observed behavior: proportion correct, RT distributions for correct responses, and RT distributions for error responses. In addition, the model provides several nontrivial predictions that have been confirmed by experiment, such as predictions about the relative speed of correct versus incorrect RTs and about the relative speed's interaction with the speed–accuracy trade-off (e.g., Ratcliff & Smith, 2004; Wagenmakers et al., 2008).

The Diffusion Model Parameters

The version of the diffusion model that we apply in this article has seven parameters:

1. *Mean drift rate* (v). Drift rate quantifies the deterministic component in the information accumulation process. This means that when the absolute value of drift rate is high, decisions are fast and accurate; thus, v indexes task difficulty or participant ability.

2. *Across-trial variability in drift rate* (η). This parameter reflects the fact that drift rate may fluctuate from one trial to the next, according to a normal distribution with mean v and standard deviation η. The η parameter allows the diffusion model to account for data in which error responses are systematically slower than correct responses (Ratcliff, 1978).

3. *Boundary separation* (a). Boundary separation quantifies response caution and modulates the speed–accuracy

trade-off: At the price of an increase in RT, participants can decrease their error rate by widening the boundary separation.

4. *Mean starting point* (z). Starting point reflects the a priori bias of a participant for one or the other response. This parameter is usually manipulated via payoff or proportion manipulations (Edwards, 1965; Wagenmakers et al., 2008; but see Diederich & Busemeyer, 2006).

5. *Across-trial variability in starting point* (s_z). This parameter reflects the fact that starting point may fluctuate from one trial to the next, according to a uniform distribution with mean z and range s_z. The s_z parameter also allows the diffusion model to account for data in which error responses are systematically faster than correct responses.

6. *Mean of the nondecision component of processing* (T_{er}). This parameter encompasses the time spent on common processes—that is, those executed irrespective of the decision process. The diffusion model assumes that the observed RT is the sum of nondecision and decision components (Luce, 1986):

$$\mathrm{RT} = \mathrm{DT} + T_{er}, \tag{3}$$

where DT denotes decision time. Therefore, nondecision time T_{er} does not affect response choice and acts solely to shift the entire RT distribution.

7. *Across-trial variability in the nondecision component of processing* (s_t). This parameter reflects the fact that nondecision time may fluctuate from one trial to the next, according to a uniform distribution with mean T_{er} and range s_t. The s_t parameter allows the model to capture RT distributions that show a relatively shallow rise in the leading edge.

Many experiments attest to the validity and specificity of the parameters of the diffusion model. For instance, Ratcliff and Rouder (1998), Voss, Rothermund, and Voss (2004), and Wagenmakers et al. (2008) have shown that easier stimuli have higher drift rates, that accuracy instructions increase boundary separation, and that unequal reward rates or presentation proportions are associated with changes in starting point. These and other experiments justify the psychological interpretation of the diffusion model parameters in terms of underlying cognitive processes and concepts.

The Diffusion Model and the Effect of Practice

In this article, we study the extent to which practice affects the parameters of the diffusion model, which will allow us to draw conclusions in terms of the processes postulated by these parameters. The characteristic speedup with practice could be captured by several parameters of the diffusion model, but the prime candidate for a parameter that captures the practice effect is drift rate, because it reflects the ease with which people process stimulus information.

Consistent with this intuition, Ratcliff et al. (2006) reported that drift rate increased with practice. However, their results also showed that boundary separation decreased (i.e., participants became less cautious with practice). Together, these changes in drift rate and boundary separation accounted for the observed changes in performance with practice. Unfortunately, Ratcliff et al.'s design allowed the practice effect to be assessed only across three or four sessions.

To apply the diffusion model across many practice blocks and on the level of individual participants, it would be necessary to collect a lot of data. Here, we used a 10,000-trial lexical decision experiment with 25 blocks of 400 trials each. Two participants were instructed to pay attention primarily to speed, and 2 other participants were instructed to pay attention primarily to accuracy.

EXPERIMENT

Method

Participants. Four undergraduate psychology students participated for course credit. All were native Dutch speakers.

Stimulus Materials and Design. In a lexical decision task, participants were presented with letter strings that had to be classified "word" (e.g., *fume*) or "nonword" (e.g., *drapa*). For this lexical decision task, we selected 200 low-frequency Dutch words, whose frequencies ranged from 0.31 to 5.48 per million (mean frequency = 3.44×10^{-6}, $SD = 1.29 \times 10^{-6}$; Baayen, Piepenbrock, & Gulikers, 1995). A set of 200 pronounceable nonwords was created by replacing a single letter in an existing Dutch word. (Vowels were replaced by vowels, and consonants by consonants. The words that were used to generate the nonwords were not used as word stimuli.) Words and nonwords were approximately matched in length. The set of 200 words and 200 nonwords was the same for all participants and in all blocks of the experiment.

Participants completed 5 blocks per day on 5 consecutive days. The 25 blocks of 400 stimuli thus constituted 10,000 trials per participant. Before each block, the same instructions were given. For the accuracy condition, Participants A1 and A2 were instructed to respond as quickly and accurately as possible. Their feedback was directed toward accurate responding. For the speed condition, Participants S1 and S2 were instructed to be fast, but still accurate. Here, feedback was directed toward fast responding.

Procedure. Stimuli were presented on a 17-in. CRT screen about 40 cm from the participant, using the Presentation software for Windows (Version 10.3). Letters were presented in lowercase font, 6 mm in height, in white on a black background. Responses were registered using a two-button response device attached to the computer's parallel port to achieve maximum timing accuracy. The experimenter was in the same room as the participants for the entire duration of the experiment.

For participants in the accuracy condition, RTs longer than 2,000 msec were followed by the feedback message TE LANGZAAM! (i.e., "too slow!"), and RTs shorter than 200 msec were followed by the feedback message TE SNEL (i.e., "too fast"). For RTs in the 200- to 2,000-msec time window, incorrect responses were followed by the feedback message FOUT (i.e., "error"), whereas correct responses did not trigger a feedback message. The duration of all feedback messages was 1,200 msec. Every trial started with a blank screen that was presented for 250 msec.

For participants in the speed condition, RTs longer than 750 msec were followed by the "too slow" feedback message. No feedback on accuracy was given. In all other respects, the speed condition was identical to the accuracy condition.

For every participant, the series of blocks was preceded by a short training block (with corresponding instructions) consisting of 15 words and 15 nonwords, none of which were also present in the 400 trials of the main experiment. The order of the stimuli was randomized before each block. The participants were given a 4-min break after each block. Each five-block session lasted approximately 60 min.

Results

In this section, we first summarize the data using descriptive measures of RT and accuracy, and then analyze the data using the diffusion model. All analyses were conducted on the level of individual participants.

Preprocessing of RT data. Lower bounds for acceptable RTs were determined by visual inspection, which revealed that 250 msec was a reasonable cutoff to eliminate *fast guesses*. RTs longer than 2,000 msec were designated as *slow outliers*. This filtering resulted in the elimination of 2, 2, 37, and 126 fast guesses for Participants A1, A2, S1, and S2, respectively; this means that, for Participant S2, only 1.26% of the data were discarded. In the entire experiment, only a single trial was classified as a slow outlier.

Descriptives: Emphasis on accuracy. Figures 2A and 2B show the effects of practice on accuracy and RT for Participants A1 and A2 (i.e., the accuracy condition). The upper panels show proportions of correct responses

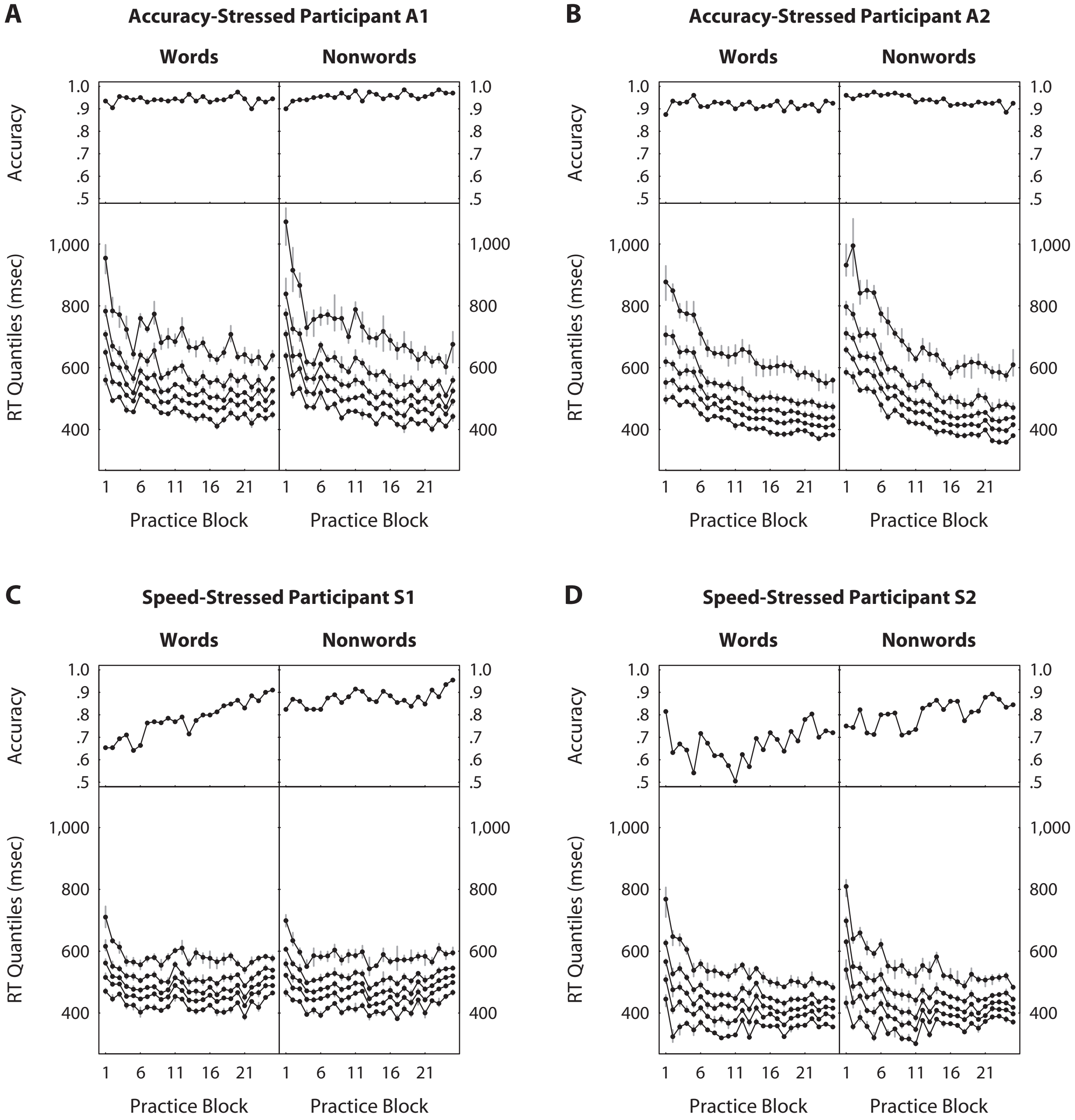

Figure 2. Mean accuracy and RT quantiles (.1, .3, .5, .7, .9) for correct responses per practice block. Gray lines in the RT quantiles show 95% bootstrap confidence intervals. Accuracy-stressed participants improved on speed. Speed-stressed participants improved mainly on accuracy. (A) Accuracy-stressed Participant A1. (B) Accuracy-stressed Participant A2. (C) Speed-stressed Participant S1. (D) Speed-stressed Participant S2.

and the lower panels show RT quantiles (.1, .3, .5, .7, and .9) calculated for each block. In addition, the figures show 95% bootstrap confidence intervals. Accuracy for words is largely constant over blocks, although accuracy for nonwords increases somewhat for Participant A1. As expected, both the mean and the spread of the RT distributions show marked decreases with practice. This pattern is evident both for correct responses (shown in Figure 2) and for error responses (shown in the supplemental materials).

Descriptives: Emphasis on speed. Figures 2C and 2D show the effects of practice on accuracy and RT for Participants S1 and S2 (i.e., the speed condition). Participant S1 clearly improves on accuracy, with constant RTs from the fourth block onward. Participant S2 appears to speed up in the first 10 blocks and then shows stable RTs with increasing accuracy. For both participants, the spread of the RT distributions decreases with practice.

In summary, accuracy-stressed Participants A1 and A2 improved mainly on speed, whereas speed-stressed Participants S1 and S2 improved mainly on accuracy. For all participants, variability in RTs decreased. Note that for all participants, performance in the final blocks was both very fast and, for all but Participant S2, very accurate. Accuracy-stressed participants seem to have started near maximal accuracy, and speed-stressed participants reached their maximum speed after a few blocks of practice.

Diffusion model analyses. We used a Bayesian parameter estimation procedure (Vandekerckhove, Tuerlinckx, & Lee, 2008) to fit the model to the data. In Bayesian estimation procedures, probability distributions quantify uncertainty about the values of the model parameters. One generally starts with a vague *prior distribution*, which then gets updated by means of the data to yield a *posterior distribution*. This posterior distribution reflects knowledge about the model parameter after having seen the data (see, e.g., Gelman, Carlin, Stern, & Rubin, 2004). We also analyzed the data with other estimation procedures (Vandekerckhove & Tuerlinckx, 2007; Voss & Voss, 2007), and this yielded results that were similar but more variable.[1]

In our estimation procedure, all parameters were allowed to vary freely across practice blocks, reflecting the exploratory nature of our analysis and the fact that we did not want to commit ourselves to a particular functional form of the practice effect. Within each practice block, drift rates were allowed to vary between words and nonwords (Wagenmakers et al., 2008), and so were the associated trial-to-trial variabilities in drift rate (i.e., the ηs); the latter modeling choice was motivated by the intuition, confirmed in early exploratory analyses, that nonwords, which by definition have no meaning, are more similar to each other than are words, which all have different meanings and frequencies. Starting point was modeled as the bias (B) in favor of words over nonwords—that is, z/a. More details about the statistical modeling and model fit can be found in the supplemental materials.

Below, we describe the modeling results, discussing in turn each of the four most important parameters (drift rate v, boundary separation a, bias B, and nondecision time T_{er}) for all participants. We only briefly mention results for the variability parameters s_z, s_t, and η; more extensive results can be found in the supplemental materials.

Diffusion model inference: Drift rate. Figures 3 and 4 show the posterior distributions of the drift rate parameter v for each of the 4 participants, for words and nonwords separately. In these figures and the ones that follow, we visualize the posterior distribution through a color coding scheme: High-density regions of the posterior have a darker color than low-density regions.

All participants but A2 show a clear increase in drift rate for both words and nonwords (Figures 3 and 4). Note that for the participants in the speed condition, drift rate increases even in the later practice blocks.

Diffusion model inference: Boundary separation. Figure 5 shows the posterior distributions of the boundary separation parameter a. Participants A1 and A2, from the accuracy condition, decrease their response caution throughout the experiment. This decrease in response caution combines with the increase in drift rate to explain why accuracy is approximately constant across practice (cf. Figures 2A and 2B), whereas RT means and variability noticeably decrease.

One might argue that, at least for Participants A1 and A2, errors are mainly caused by attentional lapses. Furthermore, one might argue that the probability of making an error due to an attentional lapse is approximately constant over practice. When RTs decrease with practice, these attentional lapses would then lead to a systematic underestimation of boundary separation. According to this account, the observed difference in boundary separation is a statistical artifact caused by model misspecification.

This misspecification account is vulnerable to at least two counterarguments. First, the supplemental materials show that both correct and error RTs decrease over time, and in a similar fashion. (The correlations of mean correct and error RTs over blocks are $r = .95$ for A1 and .92 for A2.) Second, the RT distributions for error RTs are skewed to the right, and this skew decreases with practice. These phenomena are predicted by the diffusion model, but not by the misspecification account.

For the participants in the speed condition, S1 shows little or no systematic change in boundary separation, but—just as with the participants in the accuracy condition—S2 does show a clear decrease in boundary separation with practice. Also note that, in line with the instructions, the participants in the accuracy condition have larger boundary separations (i.e., more response caution) than do the participants in the speed condition.

Diffusion model inference: Response bias. Figure 6 shows the posterior distributions of the response bias parameter B. This parameter gives the height of the starting point z as a proportion of boundary separation a, so that $B = z/a$. Thus, values of $B > .5$ indicate an a priori bias toward the word response, and values of $B < .5$ indicate an a priori bias toward the nonword response. Both participants in the accuracy condition start the experiment with a slight bias in favor of a word response and, over the practice sessions, develop a reverse preference for

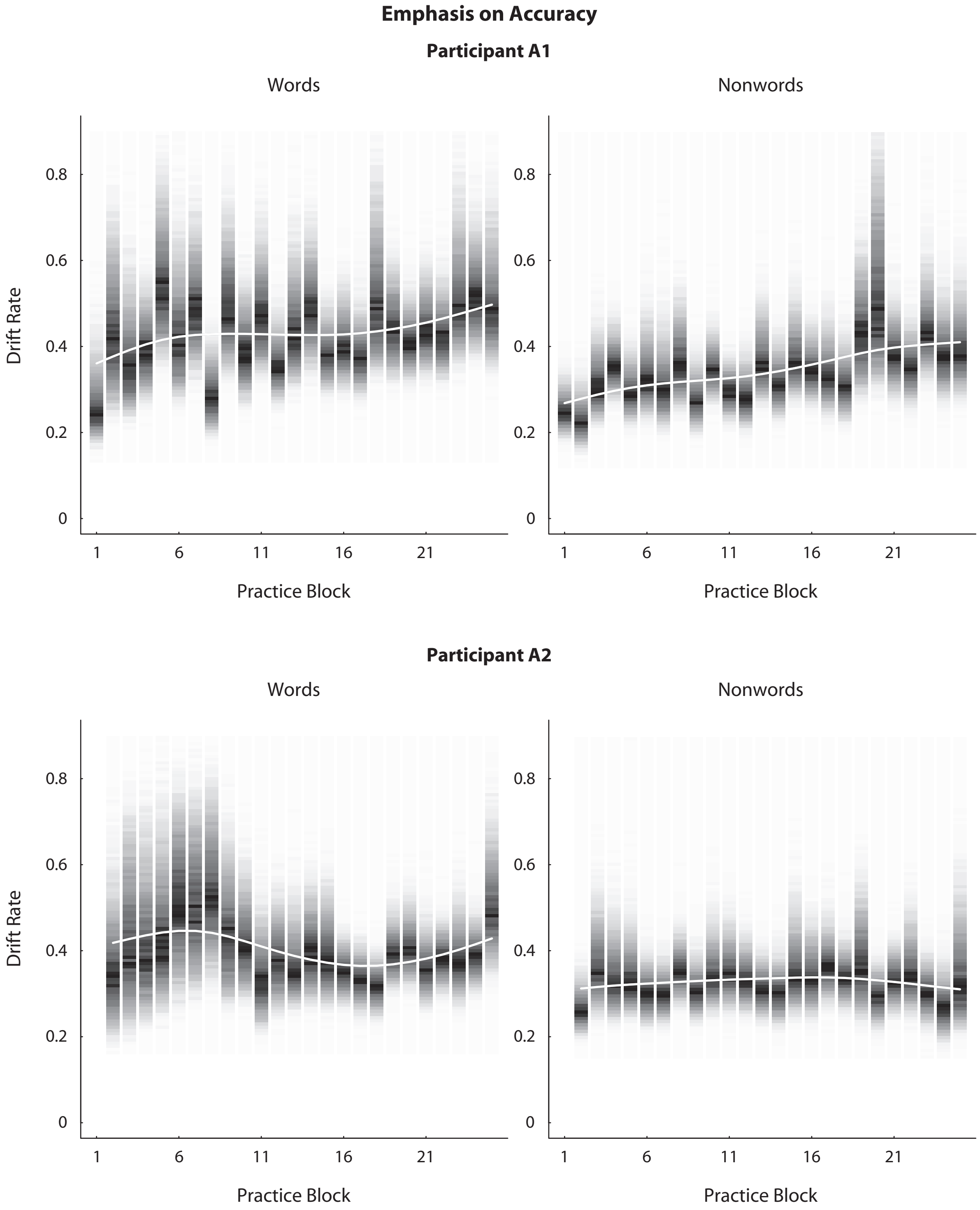

Figure 3. Posterior distributions of drift rate parameter *v* across practice blocks (for accuracy-stressed participants). Dark colors represent high density. The white lines are cubic smoothed splines through the medians of the posterior distributions.

nonword over word responses. For the participants in the speed condition, the practice-induced changes in bias are less systematic.

Diffusion model inference: Nondecision time. Figure 7 shows the posterior distributions of the nondecision time parameter T_{er}. For both participants in the accuracy condition, T_{er} decreases with practice (i.e., about 100 msec for both participants). These decreases in T_{er} account for approximately 40% of the total practice-induced decrease in mean RT, which is about 250 msec for both A1 and A2.

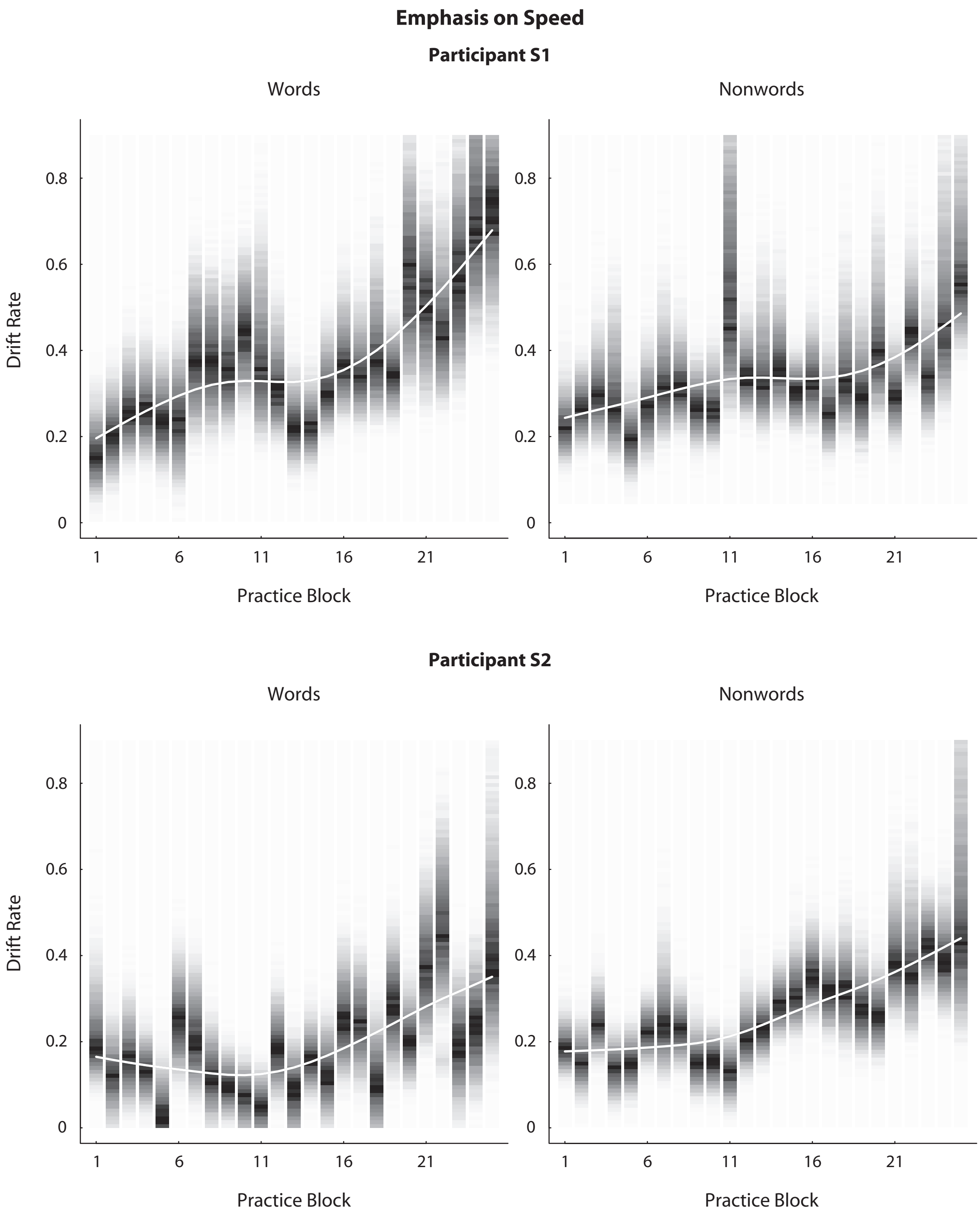

Figure 4. Posterior distributions of drift rate parameter *v* across practice blocks (for speed-stressed participants). Dark colors represent high density. The white lines are cubic smoothed splines through the medians of the posterior distributions.

The participants in the speed condition do not show a systematic decrease in T_{er} with practice, but they do display large block-to-block fluctuations in T_{er} that cover a range of about 100 msec.

Diffusion model inference: Variability parameters. In our Bayesian analyses, we estimated all parameters of the Ratcliff diffusion model, including the parameters that represent trial-to-trial variability in drift rate (η), in starting point

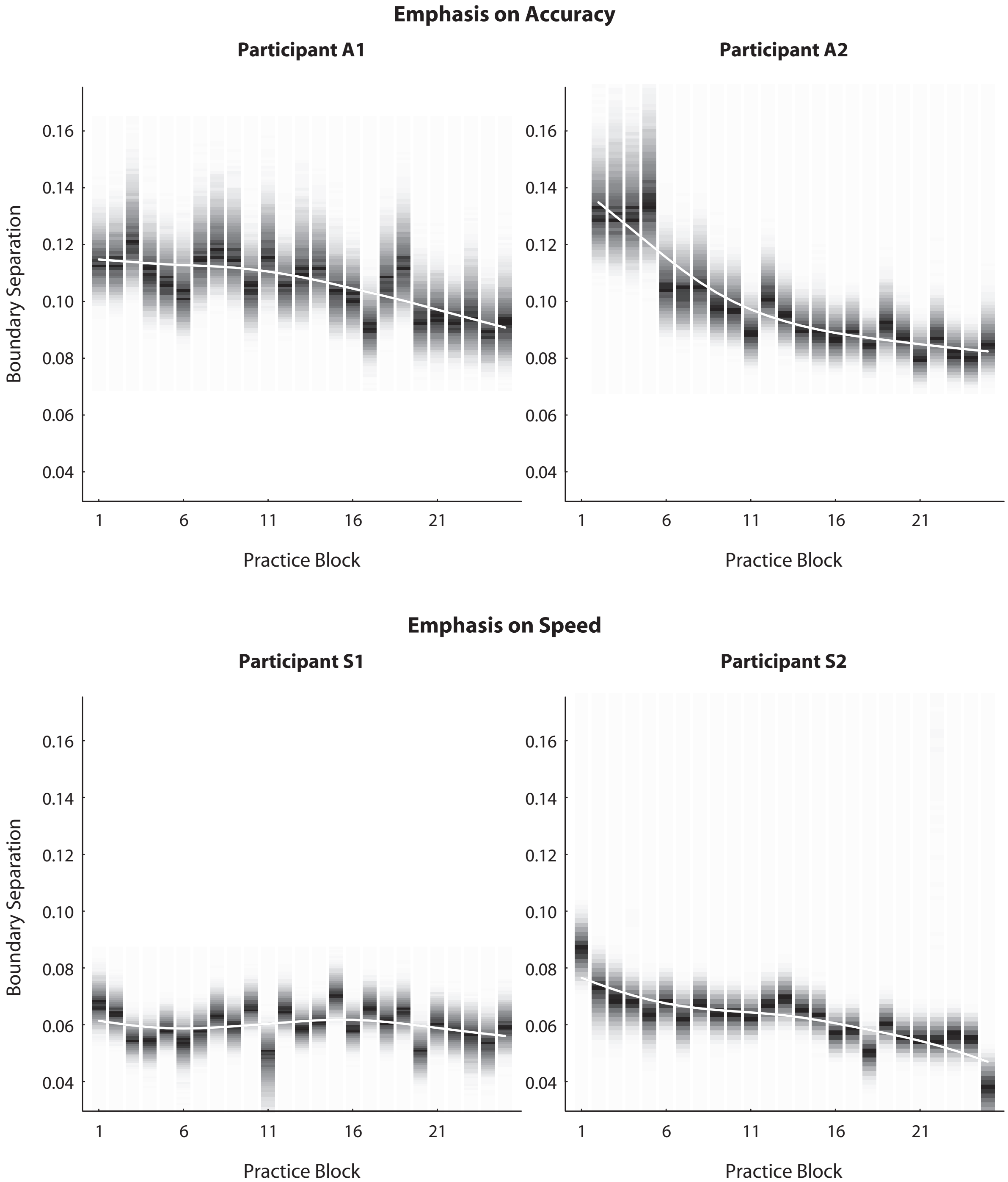

Figure 5. Posterior distributions of boundary separation parameter *a* across practice blocks. Dark colors represent high density. The white lines are cubic smoothed splines through the medians of the posterior distributions.

(s_z), and in nondecision time (s_t). The results show that, as expected, η was much higher for words than for nonwords. Also, η appeared to decrease with practice for word stimuli in the accuracy condition. The s_t parameter decreased with practice for all participants but S1. For s_z, no structural effects of practice were found. Detailed results regarding the variability parameters can be found in the supplemental materials.

CONCLUDING COMMENTS

According to our diffusion model decomposition, practice leads to an increase in the rate of information processing, a decrease in response caution, and a decrease in nondecision time. In addition, participants also exhibit systematic changes in a priori bias.

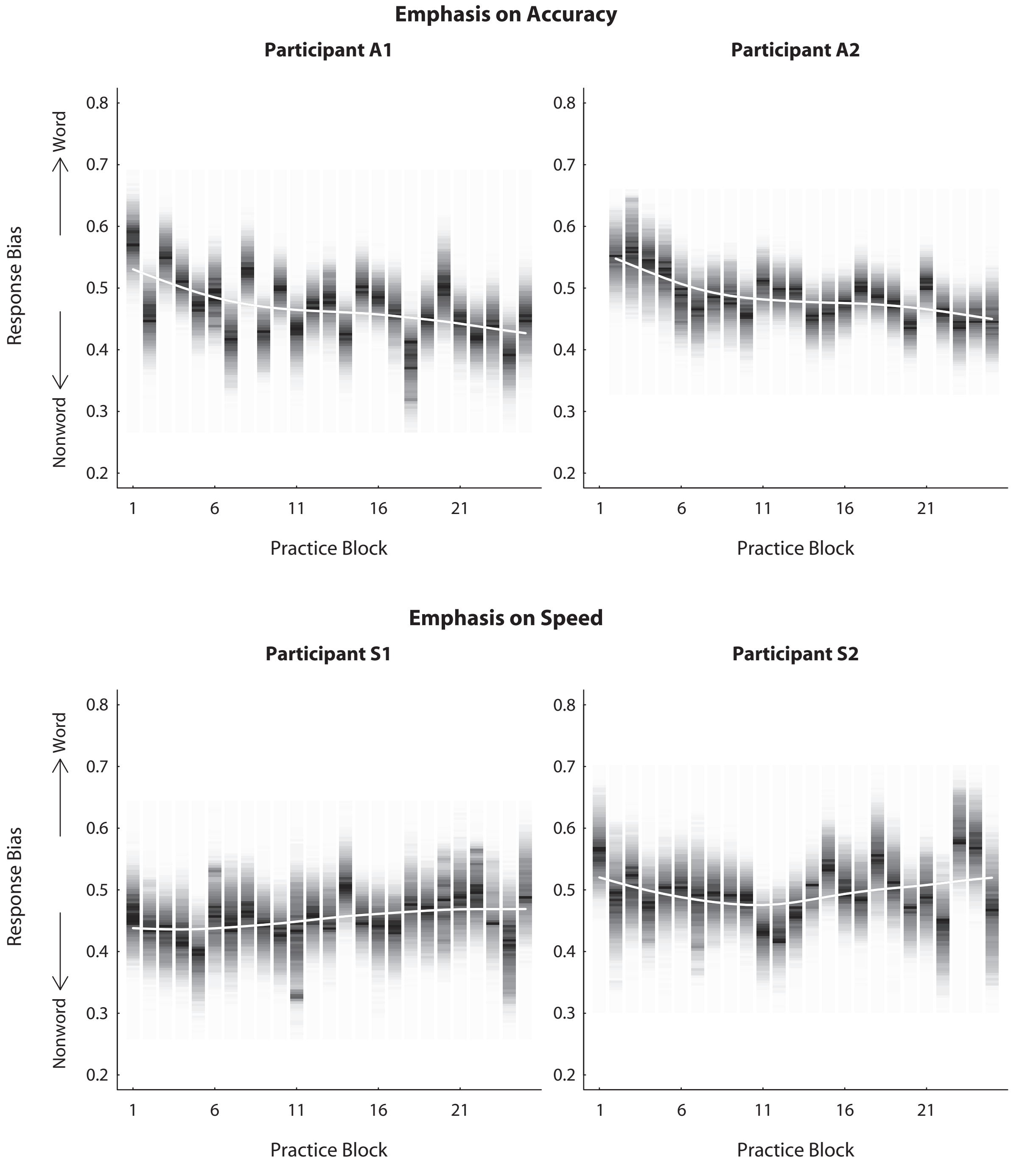

Figure 6. Posterior distributions of response bias parameter B across practice blocks. Here, B is defined as z/a. Dark colors represent high density. The white lines are cubic smoothed splines through the medians of the posterior distributions.

Among these results, the practice-induced reduction of the nondecision component and the fluctuations in response bias are both pronounced and unexpected. It is possible that the reduction in nondecision time is task- rather than stimulus-specific, hence reflecting increased familiarity with the general task requirements, the response buttons, and the processing of visual input and feedback displayed on the computer screen. To examine this possibility, future work should focus on transfer effects by including both old and new stimuli in the same task.

It should be acknowledged that the diffusion model is not an explanatory model of practice, and it does not describe how practice alters or adds memory representations. Ideally, one would like to fit our data to more substantive theories, such as Logan's instance theory (Logan, 1992), its successor ITAM (Logan, 2002), Nosofsky and

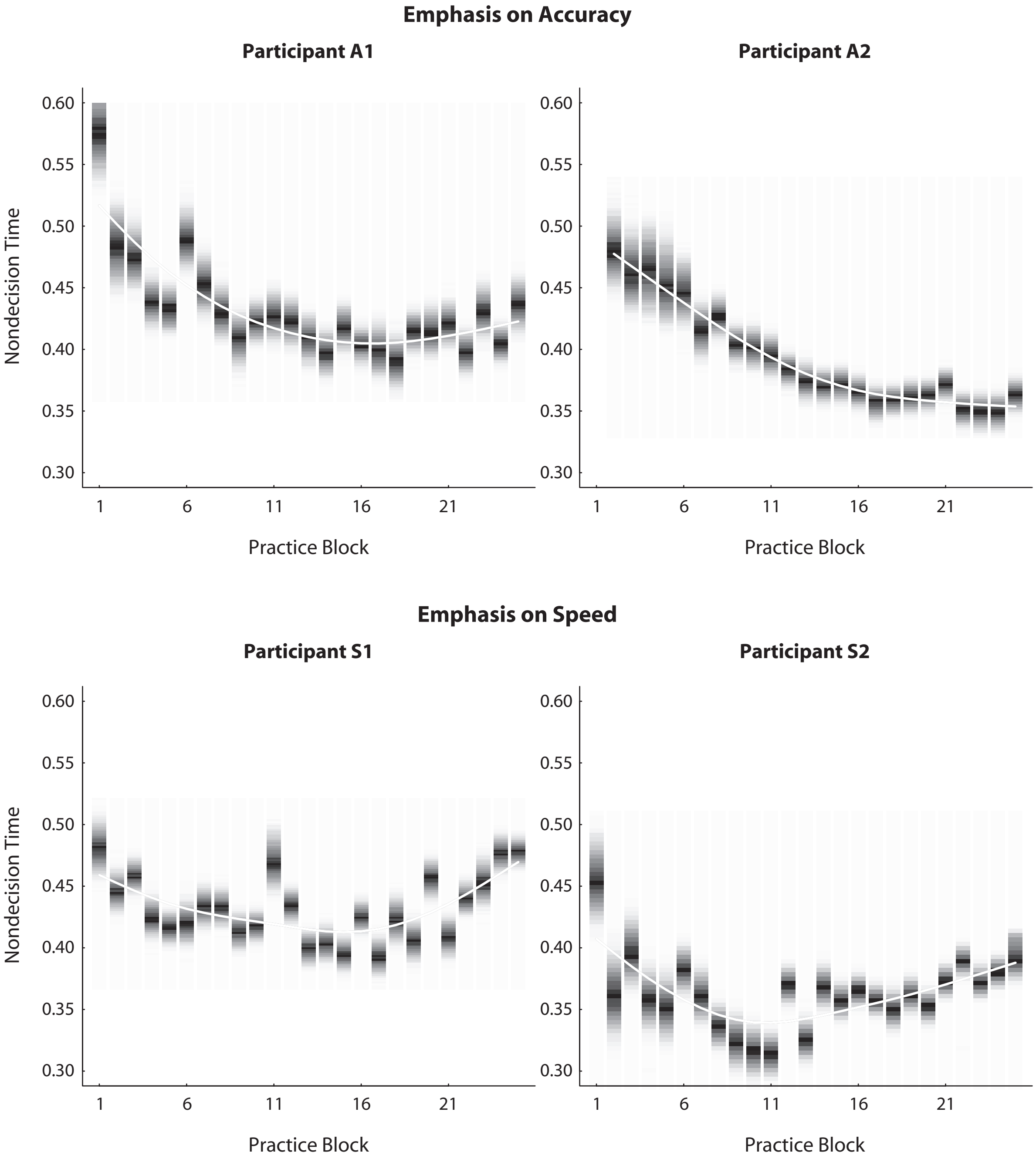

Figure 7. Posterior distributions of nondecision time parameter T_{er} across practice blocks. Dark colors represent high density. The white lines are cubic smoothed splines through the medians of the posterior distributions.

Palmeri's exemplar-based random-walk model (Nosofsky & Palmeri, 1997; Palmeri, 1997), Rickard's component power laws model (Rickard, 1997), Anderson's ACT–R (J. R. Anderson et al., 2004), or Cohen et al.'s PDP model (Cohen, Dunbar, & McClelland, 1990). Unfortunately, many of these models are less explicit about the decision process than the diffusion model. The diffusion model is able to fit entire RT distributions, both for correct and error responses, and to separately estimate components of processing such as nondecision time, response bias, and boundary separation. It is likely that the substantive models can be extended to match the performance of the diffusion model, but this currently is not the case.

In sum, the results of our diffusion model decomposition are surprising, and they suggest that the traditional methods of analysis might provide a false sense of security. Most traditional methods focus on improvements in either mean RT for correct responses or in response accuracy, without any recourse to changes in the underlying processes. Our analysis strongly suggests that the practice effect is the interactive combination of several underlying processes: People not only improve on stimulus processing, but, at the same time, are able to adjust their response strategy. In combination with changes in nondecision time, these processes generate a data pattern that cannot be usefully abstracted in terms of mean RT alone. In contrast to focusing on the mathematical function that relates practice to mean RT for correct responses (i.e., power, exponential, or APEX), we feel that a model-driven analysis of the processes underlying the practice effect will be both more appropriate and more insightful.

AUTHOR NOTE

The raw data and stimulus materials are available from the first author's Web site. We thank Marieke Jepma for her help with selecting the stimulus materials. Correspondence concerning this article may be addressed to G. Dutilh or E.-J. Wagenmakers, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands (e-mail: gilles.dutilh@gmail.com).

REFERENCES

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, **111**, 1036-1060.

Anderson, R. B., & Tweney, R. D. (1997). Artifactual power curves in forgetting. *Memory & Cognition*, **25**, 724-730.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). CELEX2 [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, **97**, 332-361.

Crossman, E. R. F. W. (1959). A theory of the acquisition of speed–skill. *Ergonomics*, **2**, 153-166.

Diederich, A., & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound-change, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics*, **68**, 194-207.

Dosher, B. A., & Lu, Z.-L. (2007). The functional form of performance improvements in perceptual learning: Learning rates and transfer. *Psychological Science*, **18**, 531-539.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology*, **2**, 312-329.

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences*, **105**, 17538-17542.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

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.

Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 883-914.

Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review*, **109**, 376-400.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

Myung, I. J., Kim, C., & Pitt, M. A. (2000). Toward an explanation of the power law artifact: Insights from response surface analysis. *Memory & Cognition*, **28**, 832-840.

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

Nosofsky, R. M., & Alfonso-Reese, L. A. (1999). Effects of similarity and practice on speeded classification response times and accuracies: Further tests of an exemplar-retrieval model. *Memory & Cognition*, **27**, 78-93.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, **104**, 266-300.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 324-354.

Pirolli, P. L., & Anderson, J. R. (1985). The role of practice in fact retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 136-153.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59-108.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, **20**, 873-922.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, **9**, 347-356.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, **111**, 333-367.

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging, practice, and perceptual tasks: A diffusion model analysis. *Psychology & Aging*, **21**, 353-371.

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.

Schouten, J. F., & Bekker, J. A. M. (1967). Reaction time and accuracy. *Acta Psychologica*, **27**, 143-153.

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*, **14**, 1011-1026.

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2008). A Bayesian approach to diffusion process models of decision-making. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1429-1434). Austin, TX: Cognitive Science Society.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, **32**, 1206-1220.

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, **39**, 767-775.

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, **21**, 641-671.

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory & Language*, **58**, 140-159.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology*. New York: Holt.

NOTE

1. Our Bayesian analysis experienced problems of numerical stability only for the first block of Participant A2. This explains why the following sections and graphs do not report any parameter estimates for this particular block of trials.

SUPPLEMENTAL MATERIALS

Further detailed information about the results discussed in this article may be downloaded from http://pbr.psychonomic-journals.org/content/supplemental.

(Manuscript received February 1, 2009;
revision accepted for publication May 29, 2009.)