

Notes and Comment

Modeling the effects of repetition and word frequency in perceptual identification

ROGER RATCLIFF and GAIL McKOON
Northwestern University, Evanston, Illinois

The counter model for word identification (Ratcliff & McKoon, 1997) has been challenged by recent empirical findings that performance on low-frequency words improves as the result of repetition of the words. We show that the model can accommodate this learning effect, and that it can do so without jeopardizing its explanations of the effects on word identification of a large number of other variables.

The counter model (Ratcliff & McKoon, 1997) was proposed as an explanation of word identification processes and the effects of word repetition on those processes. In this article, we show how the model can encompass new empirical findings that have been claimed to challenge the model (Bowers, 1999; Wagenmakers, Zeelenberg, & Raaijmakers, 2000).

The new results concern the interactive effects of word frequency and repetition in forced choice perceptual identification experiments. In these experiments, a word is flashed for a very short time, then masked, and then two alternative words are presented. Subjects are asked to decide which of the alternatives matches the flashed target. If one of the alternatives is a repetition of a word presented earlier in the experiment, subjects tend to give that word as their response, whether or not it is the correct response. In other words, the effect of prior presentation is a bias effect. The data from numerous studies (Masson, 1998; Masson & MacLeod, 1996; McKoon & Ratcliff, 1996; Ratcliff, Allbritton, & McKoon, 1997; Ratcliff & McKoon, 1996, 1997; Ratcliff, McKoon, & Verwoerd, 1989; Rouder, Ratcliff, & McKoon, 2000) showed no improvement in performance as the result of prior presentation, only a pure bias effect, and so the only effect of repetition built into the counter model was bias. However, in contrast to the earlier results, Wagenmakers et al. (2000; see also Bowers, 1999) did find that there is a small improvement in performance when low-frequency words are tested.

Given this new finding, the original claim (Ratcliff & McKoon, 1997) that the effect of prior presentation in per-

ceptual identification experiments is limited to bias is incorrect; the data from Bowers (1999) and Wagenmakers et al. (2000) show that the claim does not apply in all experimental situations. But orthogonal to this is the question of whether and how the counter model can accommodate improvements in performance for low-frequency words. It turns out that with either of two new assumptions the model provides a good account of the data. One of the possibilities is that the rate of information extraction from low-frequency words is increased from a prior presentation. The other possibility is that unfamiliarity with some unstudied low-frequency words interrupts processing so that on some proportion of trials, target information is lost and subjects have to guess. In the sections that follow, we first describe the counter model as it applies to the forced choice paradigm, and then show how it can explain improvement in processing for low-frequency words.

The Counter Model

The counter model is a variant of Morton's (1969, 1970) logogen model. The model assumes that each word in memory has a decision counter. When a target word is flashed, counts of information are accumulated, one count per unit of time, each count going to one and only one counter. In forced choice, all counts are accumulated to one or the other of the counters for the two alternatives. For the counter of a high-frequency word, the accumulation begins from a higher resting level number of counts than for the counter of a low-frequency word (as in Morton's model). The model reaches a decision when the number of counts in one of the counters exceeds the number in the other by a criterial number (k , a parameter of the model). This is a relative stopping rule of the kind used by random walk models (Laming, 1968; Link & Heath, 1975; Ratcliff, 1978, 1988; Stone, 1960), different from the absolute stopping rule used by Morton.

There are three types of counts: a diagnostic count discriminates between the two alternatives (e.g., information from the initial d for the alternatives *died* and *lied*); a non-diagnostic count is consistent with either alternative (e.g., information from the final d for *died* and *lied*); and a null count comes from random noise in the processing system. The probability that a count is diagnostic is a parameter of the model (the parameter ps). This parameter is influenced by, for example, the amount of time for which the target word is flashed: The longer the flash time, the more likely are counts to be non-null and so the more likely are diagnostic counts. Diagnostic counts are always accumulated into whichever of the two counters for the forced choice alternatives is appropriate. For nondiagnostic and null counts, each of the two counters has an equal chance (.50)

Preparation of this article was supported by NIMH Grant HD MH44640 and NIDCD Grant R01-DC01240. Correspondence should be addressed to R. Ratcliff, Department of Psychology, Northwestern University, Evanston, IL 60208 (e-mail: r-ratcliff@nwu.edu).

Table 1
Proportion of Correctly Identified Targets as a
Function of Word Frequency and Study Condition:
Experimental Data From Wagenmakers et al. (2000)

Test		Study			
Target	Foil	Target	Foil	Neither	Both
HF	LF	0.870	0.773	0.868	0.842
LF	HF	0.842	0.683	0.757	0.793
HF	HF	0.862	0.753	0.816	0.822
LF	LF	0.874	0.720	0.765	0.821

Note—HF, high frequency; LF, low frequency.

of accumulating them, if there has been prior presentation of neither alternative. With prior presentation, the counter for the repeated word becomes an attractor to nondiagnostic and null counts, stealing them away from the other counter. Stealing is modeled as an increase in the probability of the repeated word's counter accumulating null and nondiagnostic counts; in previous applications, the increase in probability, a parameter of the model, was .01 (from .50). This theft is the mechanism by which the model predicts bias. If both words, not just one, have previously been presented, the counters of both words become attractors, canceling each other out and leaving the probability that each accumulates nondiagnostic and null counts equal (i.e., at .50). Importantly, counters are assumed to be arranged in terms of similarity and the attractor force of the counter for a previously presented word is assumed to be weak, so counts can only be stolen from nearby counters—that is, from the counters of similar words. In this way, the model predicts bias from previous study for similar words but not for dissimilar words (as found by Masson, 1998; Masson & MacLeod, 1996; Ratcliff et al., 1997; Ratcliff & McKoon, 1997; Ratcliff et al., 1989; Rouder et al., 2000).

Application of the Counter Model

With the assumptions just outlined, the counter model accounted for all the data available at the time it was developed. The question is whether an added assumption can allow it to account for the data from Wagenmakers et al.'s (2000) and Bowers's (1999) new experiments. Wagenmakers et al.'s study provides a particularly good array of experimental conditions for modeling purposes. In their study, the procedure consisted of forced choice with similar alternatives. One variable was the frequencies (high or low) of the target word and the alternative word (called the *foil* by Wagenmakers et al.). The other variable was whether the flashed target, the foil, neither of them, or both of them had been previously presented in a study list that preceded the word identification tests. The conditions and Wagenmakers et al.'s data are laid out in Table 1.

For some of the conditions in Wagenmakers et al.'s (2000) experiment, the original formulation of the counter model provides accurate predictions—for example, when the target and the foil words are high frequency (the third row of the table). These data show a pure bias effect of prior study. Probability correct in the baseline condition

(neither word was studied) is .816. Study of only the target biases responses toward the target, increasing probability correct, and study of only the foil biases responses toward the foil, decreasing probability correct for the target. As predicted by the counter model, the benefit from prior study of the target (.046, the difference between .816 and .862) is about balanced by the cost from prior study of the foil (.063, the difference between .816 and .753), so that there is no significant overall improvement in performance. Also as predicted by the model, for high-frequency words, study of both the target and the foil does not change performance significantly from baseline.

The model as originally formulated also accounts for better performance with a high-frequency target than with a low-frequency target when the target and foil are not both high or both low frequency. The counter for a high-frequency target starts out at a higher resting level than that of the counter for its low-frequency alternative, but the counter for a low-frequency target starts out at a lower resting level than that of the counter for its high-frequency alternative. In contrast, when the target and the foil are both of high or both of low frequency, their two counters start out at the same resting levels.

What the assumptions of the original model cannot accommodate is the improvement in performance that comes about with prior study of low-frequency words. Wagenmakers et al.'s (2000) data clearly show this improvement. For example, in the condition in which both the target and the foil were studied, there is a significant improvement in performance of .056 over the baseline condition in which neither were studied (.821 vs. .765). Note that this improvement (as in Bowers, 1999, and McKoon & Ratcliff, in press) is smaller than the bias effect.

In the counter model, a true improvement in performance occurs when there is better extraction of information from the stimulus—that is, when the value of ps increases. In past applications, it was assumed that ps varies with flash time for the target—the longer the flash time, the more information extracted.

To model Wagenmakers et al.'s (2000) findings with low-frequency words, we examined two possibilities. The first is that without prior study, low-frequency words have a lower value of ps than high-frequency words do. This might be because low-frequency words are unfamiliar and so their features (e.g., letter combinations) are more difficult or time consuming to extract, with the result that the information input to the counters is reduced in quality. With prior study, this value of ps increases. The second possibility is that unfamiliarity with low-frequency words presented as the forced choice alternatives can sometimes, on some proportion of trials, interrupt processing of the flashed target so that no information from the target is available; in this case, subjects would have to guess which of the two choices was the correct one. In the least sophisticated version of this hypothesis, we see unfamiliarity with a low-frequency word derailing processing because a subject is trying to figure out whether the word is or is not a word, what its meaning is, and so

Table 2
Proportion of Correctly Identified Targets as a Function of Word Frequency and Study Condition: Predictions From the Counter Model

Test		Study					
Target	Foil	Target		Foil		Neither	Both
HF	LF	0.893		0.794		0.850	0.850
Parameter values		.51	+1 .08	.49	+1 .08	.50	+1 .08
LF	HF	0.834		0.674		0.748	0.779
Parameter values		.51	-1 .08	.49	-1 .07	.50	-1 .07
HF	HF	0.866		0.755		0.817	0.817
Parameter values		.51	0 .08	.49	0 .08	.50	0 .08
LF	LF	0.866		0.719		0.788	0.817
Parameter values		.51	0 .08	.49	0 .07	.50	0 .07

Note—The three parameter values shown for the model for each data point are, in order: the probability that the target attracts nondiagnostic and null counts; the resting level of the target relative to the foil; and the probability *ps* that a count is diagnostic. These are approximate values (rounded for exposition); the correct values are *ps* = .0652 for nonstudied low-frequency words, *ps* = .0752 for studied low-frequency words and for high-frequency words; and bias = .0091.

on. Of course, the interruption may not be this dramatic, but a small interruption may be sufficient to cause a loss of information from the flashed target.

With the first hypothesis, that the *ps* for low-frequency words is increased by prior study, the counter model provides a good account of Wagenmakers et al.'s (2000) data. Table 2 shows the best-fitting probability correct values for the 16 conditions. The basic equation for the model gives the probability that a count is accumulated by the flashed target's counter:

$$p = ps + (1 - ps)(.50 + bias),$$

where bias is the attractor effect of prior study.

For each condition, the table shows the values of the parameters of the model: the probability of the target attracting null and nondiagnostic counts, the resting level number of counts for the target relative to the foil, and the probability that counts are diagnostic (*ps*). The table gives rounded values of the parameters; the exact values are as follows: The attractor force of prior study increased the probability of a counter getting null and nondiagnostic counts from .5 to .5091. The difference in the resting level of high- versus low-frequency counters was +1. Without prior study, the value of *ps* for a low-frequency target was .0652. With prior study, this value increased by .010 to .0752. It was assumed that this increase brought the *ps* for low-frequency words up to equality with high-frequency words (that is, the *ps* for high-frequency words in both the studied and unstudied conditions was .0752). (This assumption of equality between studied low-frequency words and high-frequency words is not a necessary assumption, but it led to a good fit with Wagenmakers et al.'s [2000] data.) These values are all in good agreement (within 10% to 20%) of the values used for the fits of the model to all the data described by Ratcliff and McKoon (1997). The criterion number of counts by which one

counter must beat the other in order for the system to reach a decision was set at 10, the same as in the 1997 fits.

With only the three parameters free to vary—the value of *ps* for low-frequency words without prior study, the increase in this value with prior study, and the probability of a counter getting null and nondiagnostic counts as a result of prior study—the model provides a good fit to the data. The data values show a range of 19%. In only three cases does the counter model miss a data point by more than .02, and in 9 of the 16 cases, the model values are within .01 of the data values.

With the second hypothesis, that the decrement in performance for unstudied low-frequency words is caused by an interruption in processing on some proportion of trials, the counter model also fits the data well. The idea is that the presentation of low-frequency words as the forced choice alternatives can interrupt processing of information from the flashed target so that perceptual information from the target is lost. Prior study of the target makes interruption less likely, because in this case, the word has been processed relatively recently. Under this assumption, the probability of a correct response to an unstudied low-frequency word comes from a mixture of two kinds of trials: When there is no interruption, the decision process is the same process as for high-frequency words. When there is an interruption, the subject simply guesses. In other words, there is no difference in processing for high- versus low-frequency words except that on some proportion of trials for unstudied low-frequency words, target information is lost and so subjects must guess. The expression for probability correct for unstudied low-frequency words is

$$P_C(LF) = [P_C(HF) - .5] * (1 - pg) + .5,$$

where *P_C*(HF) is the probability correct (for both high- and low-frequency words) when there is no interruption

in processing, and pg is the probability of interrupted processing when one of the choice alternatives is a low-frequency word; when both choices are low frequency, pg would be $2pg - pg^2$. Fitting Wagenmakers et al.'s (2000) data, the value of ps was .075 (constant across all conditions, studied and unstudied high- and low-frequency words), the attractor force of prior study increased the probability of a counter's getting null and nondiagnostic counts from .5 to .5112, and the probability of an interruption to processing (pg) for unstudied low-frequency words was .093. With these parameters, the fit of the model to the data was about 20% better than with the assumption that ps increases for studied relative to unstudied low-frequency words. The mean absolute deviation of the model from the data was .78% with the interruption assumption, compared with .99% with the assumption that ps increases with repetition. The two assumptions handle the data about equally well (the slightly better fit with the interruption assumption is certainly not sufficient to discriminate between the two), and further work would be needed for discrimination between them.

Discussion

In the 1997 applications of the counter model to data, the only effect of prior study in perceptual identification was bias (Ratcliff & McKoon, 1997); this was consistent with all of the available empirical results. Now, with either of two added assumptions, the model can also account for Wagenmakers et al.'s (2000) and Bowers's (1999) new results that the identification of low-frequency words is improved by prior study. One of the possible assumptions is that, for low-frequency words, prior study increases the probability of extraction of diagnostic information from the flashed target; the other is that on some proportion of the trials, unfamiliarity with an unstudied low-frequency word derails processing. Wagenmakers et al. have claimed that the addition of such assumptions is a basic change in the model, but this is not the case. Rather, the change is in assumptions about the behavior of parameters of the model, not in the model itself.

It is important to recognize that, in the counter model, an improvement in processing is separate from bias. An improvement in processing, whether from increased flash duration or from prior study of a low-frequency word, lies in the extraction of information. Bias lies in the attraction of counts to a counter. This suggests the possibility of a dissociation, and McKoon and Ratcliff (in press) have demonstrated such a dissociation empirically. The intuition behind the experiment was that study of a low-frequency word would reduce the chance that it would appear novel for days (or even weeks). In contrast, the bias effect in forced choice decays over a period of hours; a delay of 48 h between study and test is sufficient for bias to largely dissipate (for bias and priming effects of the size at issue here, see Ratcliff & McKoon, 1997, Experiment 7; Whitlow & Dalton, 1999). In the McKoon and Ratcliff (in press) experiment, study lists were separated from

perceptual identification tests by either the usual delay of a few seconds or a delay of 48 h. The results of the experiment conformed to expectations: Study of a low-frequency word 48 h prior to test improved performance, but did not lead to bias; only immediately prior study gave bias. Also, studying a low-frequency word twice, once 48 h and once immediately prior to test, gave no additional improvement to performance in comparison with studying the word only once 48 h prior to test.

Separating a performance improvement from bias is a strength of the counter model relative to the older logogen model on which it is based. Models like the logogen model (Morton, 1969) attribute word-frequency effects and repetition effects to the same mechanism. The well-known problem with this is that several repetitions in an experiment of a low-frequency word should, according to the model, turn it into a high-frequency word, but, of course, this does not happen. Also, if some word is presented often enough in an experiment, the prediction is that the resting level number of counts in its counter will increase to threshold, and then, because it is always the first word to reach threshold, it will always be the word given as a response; obviously, this prediction is not supported by data. These problems necessitate additional assumptions for models like the logogen model, whereas the counter model handles the issue by assigning repetition and word-frequency effects to different mechanisms.

In sum, with the addition of either of two assumptions about the behavior of one of its parameters (but no change in the model itself), the counter model accounts for the data from forced choice perceptual identification experiments with low- as well as high-frequency words. In addition, as shown originally, it explains the data from experiments using naming and yes/no identification paradigms; it can fit the data of individual subjects across paradigms; it explains the effects of a number of variables including repetition, target duration, similar versus dissimilar alternatives in forced choice, and neighborhood size (the number of words similar to a target); and it appropriately decouples the effects of word frequency and repetition. As in 1997, our hope is that the model will be the impetus for new empirical results that will lead to a new and better model. In addition, the advent of articles that address the theoretical issues raised by explicit models like the counter model should raise the level of theoretical discussion in the research domain of implicit memory.

REFERENCES

- BOWERS, J. S. (1999). Priming is not all bias: Commentary on Ratcliff and McKoon (1997). *Psychological Review*, **106**, 582-596.
- LAMING, D. R. J. (1968). *Information theory of choice reaction time*. New York: Wiley.
- LINK, S. W., & HEATH, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, **40**, 77-105.
- MASSON, M. E. J. (1998, March). *Episodic influences on word identification*. Paper presented at the Tsukuba International Conference on Memory, Tsukuba, Japan.
- MASSON, M. E. J., & MACLEOD, C. M. (1996). Contributions of pro-

- cessing fluency to repetition effects in masked word identification. *Canadian Journal of Experimental Psychology*, **50**, 9-21.
- McKoon, G., & Ratcliff, R. (1996). Separating implicit from explicit retrieval processes in perceptual identification. *Consciousness & Cognition*, **5**, 500-511.
- McKoon, G., & Ratcliff, R. (in press). Bias in the counter model. *Psychological Review*.
- Morton, J. (1969). The interaction of information in word recognition. *Psychological Review*, **76**, 165-178.
- Morton, J. (1970). A functional model for memory. In D. A. Norman (Ed.), *Models of human memory* (pp. 203-254). New York: Academic Press.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59-108.
- Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling the accumulation of partial information. *Psychological Review*, **95**, 238-255.
- Ratcliff, R., Allbritton, D., & McKoon, G. (1997). Bias in auditory priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 143-152.
- Ratcliff, R., & McKoon, G. (1996). Bias effects in implicit memory tasks. *Journal of Experimental Psychology: General*, **125**, 403-421.
- Ratcliff, R., & McKoon, G. (1997). A counter model for implicit priming in perceptual word identification. *Psychological Review*, **104**, 319-343.
- Ratcliff, R., McKoon, G., & Verwoerd, M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 378-387.
- Rouder, J. N., Ratcliff, R., & McKoon, G. (2000). A neural network model of implicit memory for object recognition. *Psychological Science*, **11**, 13-19.
- Stone, M. (1960). Models for choice reaction time. *Psychometrika*, **25**, 251-260.
- Wagenmakers, E.-J. M., Zeelenberg, R., & Raaijmakers, J. G. W. (2000). Testing the counter model for perceptual identification: Effects of repetition priming and word frequency. *Psychonomic Bulletin & Review*, **7**, 664-669.
- Whitlow, J. W., Jr., & Dalton, P. (1999). *Priming is consistent, not persistent: Comparisons of priming effects in perceptual identification and stem completion*. Manuscript submitted for publication.

(Manuscript received October 18, 1999;
revision accepted for publication March 27, 2000.)