Bulletin of the Psychonomic Society
1978, Vol. 12 (2), 152-154

Is d_s' a suitable measure of recognition memory "strength"?

RICHARD A. CHECHILE
Tufts University, Medford, Massachusetts 02155

The signal-detection model of recognition memory is compared to the Chechile and Meyer (1976) recognition model. The Chechile and Meyer model is found to outperform the signal-detection model on three grounds. First, detection strength measures, like d_s', are sensitive to the degree of similarity of the distractor items to the target items, whereas the Chechile and Meyer storage parameter, θ_s, is not. Second, θ_s better predicts recognition speed. Third, d_s' is shown to confound the two orthogonal processes of storage and guessing.

Since the original study by Egan (Note 1), signal-detection measures such as d' and d_s' have often been used in the analysis of recognition memory data. While signal-detection theory does represent a sophisticated attempt to correct for response bias factors, there are nevertheless a number of vexing problems with the application of detection theory to recognition memory data.

First of all, a paradox results by using detection theory to scale recognition memory strength. Namely, the strength values should decrease as an increasing function of the similarity between distractor items and the target items, since strength is taken as the separation between the noise and signal-plus-noise distributions. In the perceptual case, it makes perfectly good sense to consider the discrimination of a signal to be reduced in a noisier context because the signal is embedded within the noise. However, in a yes-no rating recognition task, the target item is not embedded among distractor items. Through the use of the signal-detection model, trace strength is scaled relative to the particular class of distractor items used.

Increased distractor similarity does in fact reduce d_s' values, as will be shown with data from Gerrein (1976). In that study, each of 40 subjects had a number of recognition memory trials using the Brown-Peterson paradigm. The subjects had a yes-no recognition task followed by a three-point confidence rating with a "3" rating denoting highest confidence. The subjects had recognition testing where the distractors were either similar or dissimilar to the target items. The group data for the 12-sec retention interval are shown in Table 1. Applying detection theory to the group data yields $d_s' = 3.48$ in the dissimilar condition and $d_s' = 1.74$ in the similar condition. Thus d_s' is a measure that is scaled relative to the characteristics of the distractor distribution. However, it is not necessarily inevitable that recognition memory measures need be relativistic measures like d_s'. In fact, the same Gerrein data can be reanalyzed by an alternative recognition memory model that has a measure that is invariant with respect to distractor similarity.

More specifically, Chechile (1973) and Chechile and Meyer (1976) have presented an analysis of old and distractor recognition as well as recall trials. Although the major purpose of this appraoch is in the separation of storage and retrieval factors in memory, the model does include an analysis of old and distractor recognition tasks. A description of the relevant task analysis can be found in studies by Chechile (1973), Chechile and Butler (1975), and Chechile and Meyer (1976, Model 3). For the recognition task, there are three model parameters of primary interest: θ_s, the proportion of times that the target information has been sufficiently stored; θ_g, the proportion of times that the subject can guess correctly on old recognition when the target is not sufficiently stored; $\theta_{g'}$, the proportion of times that the subject can guess correctly on distractor recognition when the target is not sufficiently stored. Estimates of θ_s, θ_g, and $\theta_{g'}$ for the Gerrein (1976) data were determined by using the modes of the respective distributions generated by Equations 9, 11, and 12 from Chechile and Meyer (1976). Also, the standard deviation on each parameter can be computed from the previously mentioned parameter distribution equations of Chechile and Meyer (1976). The resulting point estimates and standard deviations are shown in Table 1. The outcome of the Gerrein (1976) experiment is clear; only the distractor recognition guessing parameter $\theta_{g'}$ reliably changes as a function of distractor similarity. The finding that θ_s is invariant is impressive in light of the corresponding variation of d_s' with distractor similarity.

Given the above demonstration, one might simply conclude that the two models are different and that signal-detection measures are just relativistic measures. What is needed is further empirical grounds to test and compare these competing models of recognition. Reac-

This research was supported in part by grants from the National Aeronautic and Space Administration NCA2-OR785-801 and the National Institute of Mental Health MH-12672-07. The author would like to acknowledge the able assistance of Stephanie Kalfayan in this study.

Table 1
Total Recognition Responses and the Chechile and Meyer (1976) Analysis for the Gerrein (1976) Study on Distractor Similarity

	Total Recognition Responses													Chechile and Meyer Analysis					
	Similar						Dissimilar							Storage		Old Recognition Guessing		Distractor Guessing	
	Yes			No			Yes			No									
	1	2	3	1	2	3	1	2	3	1	2	3		θ_s	SD	θ_g	SD	θ_g'	SD
Old	9	72	310	5	33	51	6	61	333	9	21	50	Similar	.630	.066	.510	.065	.510	.080
Distractor	2	36	52	7	40	343	1	0	0	0	5	474	Dissimilar	.670	.072	.490	.075	.990	.010

tion time information could be one way to compare the suitability of these two models of recognition memory. It is reasonable to expect shorter recognition times for "stronger" or sufficiently stored items as compared to "weaker" items. Thus reaction time is ideally suited to be a model-free criterion variable. Consequently, the main purpose of the present study is to use reaction time data to evaluate and compare the two models of recognition memory.

METHOD

Subjects

Thirty undergraduate students at Tufts University participated in the experiment.

Materials and Apparatus

There were 228 consonant trigrams that were used in recognition memory testing. The stimuli were 152 trigrams randomly selected from the meaningfulness range of 13% to 22% in the Witmer (1935) norms. Seventy-six additional trigrams were used as distractors.

All trials were recorded on one track of a tape. On the second track, a signal was recorded at the start of the test cue. This signal activated a Uher 420 diapilot, which in turn activated a Gerbrands digital millisecond counter. The timer stopped when the subjects pressed a response button. The tape was played on a TEAC 2300S recorder and was heard over Koss Pro 4 AA headphones.

Design and Procedure

The 152 recognition trials were subdivided into 76 randomly presented old recognition trials and 76 distractor recognition trials. Each trial consisted of three stages: the presentation of the target trigram, an interpolated interval, and a test cue. A buzzer ready signal preceded the presentation of the target trigram by .5 sec. The trigrams were presented in a female voice at a rate of 1 letter/.5 sec. Immediately after the presentation of the target, 10 digits were read in a male voice at the rate of 1 digit/.5 sec. The subjects were required to shadow the digits. After 5 sec of interpolated activity, the old or distractor trigram was presented in a female voice for 1.5 sec. The subjects were required to press the reaction time button as soon as they knew their response. The subjects were required to answer yes or no and also to give a confidence rating. The subjects were instructed to give a "3" rating if they were absolutely sure, a "2" rating if they gave an "educated" guess based on partial information, and a "1" rating if they just randomly guessed. The subjects were given 4 sec for recognition testing followed by a 2-sec intertrial rest interval. Each subject received five practice trials in the same manner before the beginning of the experimental session.

RESULTS AND DISCUSSION

The median reaction time (RT) on old recognition was determined for each subject. Across subjects the median RT measure was a skewed distribution. Consequently, a reciprocal square-root transformation was employed to achieve a symmetrical distribution. Henceforth, this transformed measure will be called speed (s). Furthemore, since many of the linearized ROC curves did not have a slope of one, the generalized signal-detection measure d_s' was computed for each subject (cf. Green & Swets, 1966, p. 98). Moreover, the parameters θ_s, θ_g, and $\theta_{g'}$ were computed on an individual subject basis by means of Model 3 from Chechile and Meyers (1976), as described earlier in the introduction. Finally, for simplicity, the two guessing measures were averaged to form a new guessing measure that will be called θ_{g*}.

The summary statistics for s, d_s', θ_s, and θ_{g*} are displayed in Table 2. The diagonal terms are the sum of the squared deviations on the measure. The above-diagonal terms are the sum of the deviation cross-products between the row and column measures. The below-diagonal terms are the product-moment correlations between the row and column measures. The means for each measure are shown in the last row.

The correlation between speed and θ_s was .700, whereas the correlation between speed and d_s' was only .402. Moreover, scatter plots for both of these correlations were examined for evidence of curvature that might account for a lowering of either correlation. The scatter plots showed that neither parameter could markedly improve its respective correlation to s by a nonlinear transformation. Also a multiple stepwise

Table 2
Means, Correlations, Deviation Sum of Square, and Deviation Sum of Cross-Products for s, d_s', θ_s, and θ_g*

	s	d_s'	θ_s	θ_g*
s	.082	.846	.161	.021
d_s'	.402	54.455	3.203	1.315
θ_s	.700	.537	.652	−.024
θ_g*	.145	.355	−.058	.252
Mean	.759	3.723	.677	.791

regression analysis was performed. If θ_s is entered first and d_s' second, then the square of the coefficient of multiple regression is .491, which is not significantly different from the squared correlation coefficient between θ_s and s [$F(1,27) = .04$]. Consequently, the inclusion of d_s' in the multiple regression does not significantly increase the predictability of the speed scores from that already accounted from the entering of θ_s first. However, the inclusion of θ_s in the regression analysis does significantly increase the predictability of the speed scores from that obtained by entering d_s' first [$F(1,27) = 17.519$, $p < .001$]. Thus the storage measure, θ_s, can account for all of the variability in speed that d_s' can predict, plus an additional amount of variability in speed that is not accountable by d_s'. Clearly, the Chechile and Meyer (1976) model of recognition memory performs better than the signal-detection model in regard to predicting speed scores.

Finally, a regression analysis was performed to investigate the relationship between the d_s' scores and the parameters θ_s and θ_{g*}. In this analysis the d_s' scores were used as the criterion variable whereas θ_s and θ_{g*} were used as predictor variables. Actually, θ_s and θ_{g*} are ideally suited as predictor variables since the correlation between θ_s and θ_{g*} was low and nonsignificant. The stepwise regression analysis showed that if θ_s is entered first, then the addition of θ_{g*} significantly improves the prediction of d_s' [$F(1,27) = 7.19$, $p < .025$]. Moreover, if θ_{g*} is entered first, then the addition of θ_s also significantly contributes to predicting d_s' [$F(1,27) = 15.02$, $p < .001$]. Consequently, both predictor variables should be included in the multiple regression analysis. The resulting regression equation is: $(d_s')_p = 5.116\,\theta_s + 5.693\,\theta_{g*} - 4.243$, where $(d_s')_p$ is the predicted d_s' score. The coefficient of multiple regression is .662. Thus d_s' has been additively decomposed into two noncorrelated components, θ_s and θ_{g*}. The importance of this finding is that studies using d_s' as the dependent variable of recognition memory strength have confounded the two orthogonal processes of storage and guessing. Hence characterizations of d_s' over conditions where guessing processes may have changed cannot be simply interpreted as the consequence of purely a storage process. This finding thus adds further support for the preference of the Chechile and Meyer (1976) model of recognition memory over the signal-detection model.

REFERENCE NOTE

1. Egan, J. P. *Recognition memory and the operating characteristic.* Technical Note AFCRC-TN-58-51, Hearing and Communication Laboratory, Indiana University, 1958.

REFERENCES

Chechile, R. A. *The relative storage and retrieval losses in short-term memory as a function of the similarity and amount of information processing in the interpolated task.* Unpublished doctoral dissertation, University of Pittsburgh, 1973.

Chechile, R., & Butler, K. Storage and retrieval changes that occur in the development and release of PI. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 430-437.

Chechile, R., & Meyer, D. L. A Bayesian procedure for separately estimating storage and retrieval components of forgetting. *Journal of Mathematical Psychology*, 1976, **13**, 269-295.

Gerrein, J. R. *Organizational, storage, and retrieval factors in alcohol-induced forgetting.* Unpublished doctoral dissertation, Tufts University, 1976.

Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics.* New York: Wiley, 1966.

Witmer, L. R. The association value of three-place consonant syllables. *Journal of Genetic Psychology*, 1935, **47**, 337-359.

(Received for publication May 9, 1978.)