美作女子大学・美作女子大学短期大学部紀要
1997, Vol. 42, 11～25

論　文

Components in "Same"-"Different" Judgments as an Interface between Perception and Higher Cognition (2)

知覚とより高次の認知をつなぐインターフェイスとしての同異判断（2）

Masahiko Saito

Saito(1996) proposed a theory of "same"-"different" judgment reaction time. The purposes of the theory were as follows: (1) to explain matching phenomena; (2) to provide links between them and our knowledge regarding perception and cognition; and (3) to demonstrate a processing principle: The principle states that in pattern matching judgments, reaction times and error rates do not depend on quantitative judgment-criteria settings but on decision rules across several description viewpoints and on differing rules across resolution levels of analysis. The present research aims at mathematical modeling of a subset of the theory. Predictions derived from the mathematical model fit observed mean reaction times, though the model is only a first approximation of the theory; predicted standard deviations of reaction time are too large[1] as compared with observed values. Because it is assumed that each probability distribution of elementary-processing times is exponential, the total theoretical system has rather large standard deviations of reaction time. Under this assumption, however, we can easily obtain values of prediction, and the choice of the probability distribution function was based on a convenience of mathematical calculation; if some other function was assumed, mathematical derivations would be too complex to find solutions. The present model, hence, is only a first approximation; but the predictions of reaction-time means and error rates will clarify, more than a little, characteristics of the present theory.

In the theory, a descriptor has four subprocesses; Level 1 outputs very global descriptions of a visual stimuli, and Level 4 describes very fine details of them. Each of the subprocesses produces multiple descriptions of a stimulus; every descriptions produced by a subprocess are identical-resolution ones, but descriptions are based on different viewpoints; and a different set of symbolic units is used for producing each different description. That is, each individual subprocess analyzing a stimulus at a level of resolution produces a set of multiple descriptions which are based on various viewpoints. In such subprocesses, "same"-"different" judgment results from a search for a correspondence between descriptions of a stimulus pair: Each of four different-resolution subprocesses of the descriptor produces two sets of descriptions, one set for each of two stimuli. When "same"-"different" judgment is needed, it searches across the two sets. Each individual subprocess then cries "same !" when it

---

This research was supported by Grant-in-Aid for Encouragement of Young Scientists (121-8002-61710102 in 1986 and 121-8003-62710098 in 1987) from The Ministry of Education, Science and Culture of Japan.

The author gratefully acknowledges the advises by Naoya Mori and by Daynise Irene Holloway.

finds a correspondence between descriptions. All descriptions produced by a given subprocess are the result of identical-resolution analyses, and any-viewpoint match is acceptable as that-level "sameness" (though different viewpoint descriptions help identification by a categorizer); this build-in strategy is based on Krueger's theorem (see Krueger,1978; Saito,1996). The criterion value of judgment is assumed to be constant in most predictions.

Level 3 identifies the roles main lines in each letter have (see Saito,1996). Identifying letters or "same"-"different" judgment for letter pairs is based on this-level descriptions, since such a judgment does not need the descriptions which Level 4 produces. Level-4 descriptions result from sharp-resolution analyses and express not only the main structure of the letter but tiny ornaments or slight skew as important attributes. This property is termed overquality. On the other hand, discrimination between similar but different fonts needs the Level-4 descriptions. The categorizer hence selects a level of "sameness," according to a specific purpose.

The "Fast-same" Phenomenon: A Mathematical Model of Level 3.

When the two stimulus patterns are the "same," two identical-viewpoint descriptions, each of which denotes each stimulus, correspond, providing describing errors do not exist. That is, the number of correspondences between descriptions equals the number of procedures which describe each stimulus. When the physical "same"-"different" judgment is required, correspondences between descriptions are checked. Since the correspondence checks proceed in parallel, the time course of this search process is as follows:

Because of internal noise, the probability of spurious mismatch is greater than that of spurious match (see Krueger,1978): When noise causes describing errors, most identical-viewpoint descriptions of two "same" stimuli mismatch. The following reasons can be cited: Even when noise changes any small part of a given description, the descriptions mismatch. In addition, it is very rare that by internal noise, both descriptions are changed but become the same in the end result. On the other hand, it is very rare that descriptions of two "different" stimuli match by chance. This spurious match occur only when every originally-differing part becomes identical by chance.

Accordingly the present theory assumes that each individual subprocess sends out a "different" message only when all different-viewpoint descriptions produced by it are different. On the other hand, a "same" message is sent out when any pair of identical descriptions is found. Thereby, "different" judgments by the individual subprocesses are the results of exhaustive processing, while "same" judgments result from self-terminating search. "Sameness" judgment by an individual subprocess will thus be faster than that of "difference." (Note that each component "same" judgment by individual "describing" procedures should be slower than that of "different" ones, because "sameness" demands a confirmation that all description units are individually the same. Hence, when discriminability is too great, overt "same" responses will be slower than "different" ones; see the mathematical model mentioned below.)

One factor, however, should be considered before predicting overt response. In ordinary matching, "sameness" is a unique concept, but "difference" judgment is not. For example, if pairs of identical letters are defined as "same," Level-3 "sameness" is reliable, but Level-1 or Level-2 "sameness" is unreliable. If pairs of letters having circleness are "same," only Level-2 "sameness" is reliable (see Table

1 in Saito,1996). On the other hand, even if only Level-3 "sameness" is relevant, all "differences" of Levels 1, 2, and 3 are acceptable. The categorizer must hence select a given level for "same" response; but it can accept any resolution level "difference" which is identical to or lower than the "sameness" level. This asymmetry should interact with some experimental variables. For example, if a "different" pair set includes different-sized letter pairs, Level 1 can find this difference. At the same time, if all "same" pairs consist of the same letters, the categorizer must select the "sameness" of Level 3. Since Level 1 is faster than Level 3, a number of overt "different" responses will be faster than "same" ones. The mean overt "different"-reaction-time value is an expression of various-level-"difference" detection latencies. Consequently, when both Level 1 and Level 2 cannot detect differences, the "fast-same" disparity in mean reaction time will be larger. In such a case, an overt "same"-"different" disparity in mean is a direct expression of the "fast sameness judgment" under Level 3. That is, when variability in the degree of difference for various "different" pairs (heterogeneity of stimulus difference) is large, the "fast-same" disparity in mean reaction time will be small or negative. This effect of the heterogeneity will be discussed in the next paper.

However, in letter matching, it seems that the effect of heterogeneity is not so large. Proctor and Rao(1983) did not find the effect, though Krueger(1986) found it when using carefully selected letters. These results indicate that the alphabet set they used has a fairly small degree of the heterogeneity.

A mathematical model of Level 3 was constructed to predict the reaction times and error rates of multiletter matching. This mathematical model ignores other subprocesses' performance. That is, the effect of heterogeneity has been neglected. However, since the effect is fairly small in the letter matching task, predictions will fit observed data. The following mathematical model predicts data in experiments containing both the simultaneous- and successive-presentation conditions. In these experiments, the number of differing letters on "different" strings is one. (See below for the reason for this restriction in the model description; in reality, the model can also be used for other types of "different" pairs.)

Stimulus "describing" time. It is assumed that the probability distribution function(S) of an individual "describing" latency (in an individual procedure) is

$$S(t) = 1 - \exp(-kt), \qquad (1)$$

where t represents time, and k stands for a time constant. The present model assumes the value of k is identical in every "describing" procedure.

Eriksen and Schultz (1978) concluded that in letter recognition tasks, the masking effect ceases at a stimulus onset asynchrony (SOA) of 100 msec. In the letter recognition task, the categorizer selects the description produced by the most proper procedure. It is hence sufficient to consider only one "describing" procedure in such a task. Then, it is assumed that

$$.95 = 1 - \exp(-100\,k); \qquad (2)$$

this means that the probability that "describing" processing terminates at a time less than or equal to 100 msec is 0.95. Of course, this value depends on some experimental manipulations, i.e., luminance of stimulus field, contrast between white and black regions, experimental room illumination, or so forth. However, an extreme luminance value would not be employed in ordinary matching experiments. Since the statistics employs the value of .95 as a critical point, the present model adopts this value in order to coordinate this model with the results of ordinary experiments. We thus obtain the estimated value of k from Equation 2:

$$k = .0299573.$$

Simultaneous presentation. In matching experiments, several "describing" procedures are performed. Here, the present model assumes that the number of procedures is 7, the magical number (Miller,1956).

Under the simultaneous-presentation condition, seven "describing" procedures are applied, and two sets of seven descriptions are produced (each set denotes a letter string). When two identical-viewpoint "describing" actions terminate, the descriptions are compared. The probability density function (r) of this termination latency (termination of two parallel "describing" actions) is

$$r(t) = 2\, S(t)\ S'(t), \tag{3}$$

where S' is the differential of S . That is, S equals the probability that a "describing" action for one of two stimuli has terminated at a time less than or equal to t; and that the probability density of another termination at that time is multiplied. The factor of 2 represents two exclusive events: (1) one of two actions terminates first; and (2) the other action terminates first.

After terminating the two "describing" actions, the two descriptions are compared. The present model assumes that the probability density function of this comparison latency is

$$q(t) = a \exp(- a\, t) \tag{4}$$

when a is the function of both complexity and discriminability (see below). The inverse of a is the expected value of each comparison latency. For "same" pair descriptions, this value is referred to as a_s, and for "different" pair, as a_d. Then, the probability distribution function (F) for latency when an individual set of describing and comparing terminates is

$$F(t) = \int_0^t \int_0^{t-y} r(x)\ q(y)\ dx\ dy \tag{5}$$

(see Appendix A)

(A) "Same" reaction time. Since the seven "describing"-"comparing" procedures are executed in parallel, the probability distribution function of overt "same" reaction time (base time is not included) is

$$G(t) = \int_0^t 7\{1 - F(x)\}^{7\text{-}1}\ f(x)\ dx\,, \tag{6}$$

where $f(x) = dF(x)/dx$. This means that when any one description match is found, a "same" response is elicited (self-terminating processing). This equation represents that at the time t , six of the seven matches have not yet been found. The factor 7 signifies that the first-found match will be any one of the seven matches.

Then, the expected value of overt "same" reaction time is

$$E[RT_s] = \int_0^\infty t\, dG(t)\, /\, dt\ \ dt + b\,, \tag{7}$$

when b represents the base time which is independent of G , and RT_s stands for "same" reaction time.

(B) "Different" reaction time. When all descriptions mismatch, a "different" response is elicited. In this prediction of multiletter matching, it is assumed that the categorizer always selects a "different" message from Level 3 only. In this case, the probability distribution function of the reaction time (base time is not included) is

$$H(t) = \int_0^t 7\ F^6(x)\ f(x)\, dx\,; \tag{8}$$

its expected value is

$$E[RT_d] = \int_0^{\infty} 7\ t\ dH(t)\ /\ dt\ \ dt + b\ , \qquad (9)$$

when RT_d represents "different" reaction time.

(c) Compactness and discriminability.

The value of a in Equation 4 should be a function of both compactness and discriminability. The present model assumes that each letter employed in the experiments has an identical degree of complexity and that the letter strings do not include any familiar pattern (e.g., spelling pattern). It is, then, sufficient to consider only letter-string length in calculating compactness.

Furthermore, the present model assumes that a parameter depending on discriminability is constant when the number of mismatch letter positions within a string pair is always one. Generally, when the number is constant, the value of the parameter should also be constant. Then a is defined as

$$a(L) = d\ \ C(L) \qquad (10)$$

where L represents letter-string length, C is a function depending on the compactness of the strings, and d is a parameter depending on the discriminability between one-letter stimuli. If the descriptions are the same, d is d_s, and a (L) is a_s (L); if not so, d_d and a_d (L).

Based on the present theory itself, the function C meeting the following requirements was selected: (1) C represents a monotonic decreasing function of L. That is, each expected value of individual comparison latencies is stated as $1/a$, and smaller compactness (greater length) should increase the value of $1/a$. (2) In this subprocess, the limit for describable letter-string length (L) is 4; when the length is 5, the value of a should be zero. That is, if L is a continuous number, the expected value of the comparison latency ($1/a$) becomes the infinite as L moves closer and closer to 5. (3) When L equals 1, the value of a is equal to d, since d represents the parameter which determines the expected value of one-letter pair comparison latency. (4) Even when L is 0, the value of a is clearly defined (a positive value). Thus, even when no letter is presented, some judgment is possible (two white stimulus fields are "same"). (5) The value of C depends on a power function when the value of the exponent is 1/3; this requirement is based on Stevens' law (Stevens, 1957). The Stevens' law applies to many psychophysical functions (loudness, brightness, numerousness, heaviness, and so forth; Stevens, 1957). The exponent value, 1/3, is adopted as a first approximation. It seems that this value is a candidate for psychological constants in visual information processing. In judging of brightness, the value is about . 3 (Stevens, 1957, p. 168). In addition, Ueno(1976) employed the reaction time measure and found that the value was . 34 or . 36 for single light pulses, while for double light pulses it was above . 27 and below . 38 when the temporal summation of the two pulses was obtained. Therefore two different types of measure indicate the value of 1/3.

Accordingly, C is assumed to be as follows:

$$C(L) = (1/2 - 1/6)^{-1/3}\ \{1/(L+1) - 1/6\}^{1/3} \qquad (11)$$

If the number of mismatch positions is two, the term of $(1/2\text{-}1/6)^{-1/3}$ should be changed to $(1/3\text{-}1/6)^{-1/3}$. In such a case, d represents the time constant of an individual comparison for two-letter-string pairs. Note that when the value of d_d is too large, overt "same" response will be slower than "different" response. For example, when the number of differing letters on strings is too large, this phenomena will occur. It should be noted that when L equals 0, Equation 10 ($a(L)$) cannot be used to calculate reaction times, although C defines compactness for blank stimuli (L =0). Equation 10 involves d which relates to one-letter comparisons, and calculation of reaction times for blank stimuli requires another constant other than d.

From Equations 2, 7, 9, 10, and 11, we obtain the set of equations as follows:

$$E\{RT_s(L)\} = \sum_{h+i+j=7} \frac{7!}{h!\ i!\ j!} \left(\frac{2a_s(L)}{a_s(L)-k}\right)^h \times \left(-\frac{a_s(L)}{a_s(L)-2k}\right)^i \left[\frac{2k^2}{\{a_s(L)-k\}\{a_s(L)-2k\}}\right]^j \times \frac{1}{hk+2ik+ja_s(L)} + b \quad (12)$$

(For the expected value of "same" reaction time occurring under the simultaneous-presentation condition, see Appendix B).

$$E\{RT_d(L)\} = -\sum_{\substack{g+h+i+j=7 \\ (g \neq 7)}} \frac{7!}{g!\ h!\ i!\ j!} \left(-\frac{2a_d(L)}{a_d(L)-k}\right)^h \times \left(\frac{a_d(L)}{a_d(L)-2k}\right)^i \left[-\frac{2k^2}{\{a_d(L)-k\}\{a_d(L)-2k\}}\right]^j \times \frac{1}{hk+2ik+ja_d(L)} + b \quad (13)$$

(For the expected value of "different" reaction time under the simultaneous-presentation condition: see Appendix C).

$$k = .0299573 \quad (14)$$

(The time constant of each "describing" action).

$$a_s(L) = (1/2 - 1/6)^{-1/3}\ d_s\{1/(L+1) - 1/6\}^{1/3}; \quad (15)$$

$$a_d(L) = (1/2 - 1/6)^{-1/3}\ d_d\{1/(L+1) - 1/6\}^{1/3}; \quad (16)$$

(The inverse of expected value of each comparison latency as a function of letter-string length, L).

Except for the base time, the number of free parameters is one in each individual response type prediction; either d_s or d_d. Because these parameters are the constant of proportionality, their different values do not produce different function shapes in the expected values. The processing algorithm and the function, $C(L)$, determine the shape of the expected-value functions, and the function, $C(L)$, includes no free parameter. Accordingly, the present mathematical model has fairly strong refutability; hence, its predictions have empirical significance (cf. , Proctor, 1986).

Successive presentation. Under the successive-presentation condition, when a test string is presented, one set of descriptions for a first-presented string is already stored in a working memory (e.g., Baddeley & Hitch, 1974). In addition, description comparisons are affected by the priming which is produced by the processing of the first-presented string.

Because one set of descriptions is in memory,

$$r^*(t) = k \exp(-k\,t) \quad (17)$$

where r^* represents the probability density function of an individual-viewpoint "describing"-action termination; then,

$$F^*(t) = \int_0^t \int_0^{t-y} r^*(x)\ q(y)\ dx\ dy \quad (18)$$

and F^* represents the probability distribution function of the latency when an individual processing set, describing and comparing, terminates under the successive-presentation condition.

Substituting Equation 18 for F in Equation 7 or in Equation 9, we obtain either the expected value of "same" reaction time or of "different" one, respectively (see Appendix D).

Thus we obtain the following set of equations:

$$E\{RT_{*s}(L)\} = \int_0^\infty t\, 7[1 - F^*(t)]^6 f^*(t)\; dt + b$$

$$= \sum_{i=1}^{7} \frac{7!}{i!\,(7-i)!} \left(\frac{k}{k - a_{*s}(L)}\right)^i \left(-\frac{a_{*s}(L)}{k - a_{*s}(L)}\right)^{7-i}$$

$$\times \frac{1}{i\, a_{*s}(L) + (7-i)\,k} + b \qquad (19)$$

(The expected value of "same" reaction time; a_* represents the inverse of expected value of each comparison latency under the successive-presentation condition: see below).

$$E\{RT_{*d}(L)\} = \int_0^\infty t\; 7[F^*(t)]^6 f^*(t)\; dt + b$$

$$= -\sum_{\substack{h+i+j=7 \\ (h \neq 7)}} \frac{7!}{h!\; i!\; j!} \left(-\frac{k}{k - a_{*d}(L)}\right)^i$$

$$\times \left(\frac{a_{*d}(L)}{k - a_{*d}(L)}\right)^j \frac{1}{i\, a_{*d}(L) + jk} + b \qquad (20)$$

(The expected value of "different" reaction time under the successive-presentation condition).

Under this condition, priming should affect each comparison; it is assumed that the priming effect (T) is

$$T(L) = u\,(L+1)^{1/3}, \qquad (21)$$

where u is the priming constant. The term of L+1 means that even when no letter is presented, priming for nothing occurs. The inverse expected values of the comparison latencies (a_*) are

$$a_{*s}(L) = d_s\;\; C(L)\;\; u_s\;\; (L+1)^{1/3} \qquad (22)$$

$$a_{*d}(L) = d_d\;\; C(L)\;\; u_d\;\; (L+1)^{1/3} \qquad (23)$$

when u_s and u_d are the priming constants in individual "same" and "different" judgments, respectively.

The value of k is equal to that of Equation 14. Here, the values of both d_s and d_d must equal the values estimated under the simultaneous-presentation condition, since an identical stimulus set is used under both conditions. In addition, the base time value should also be identical. If these values are estimated from reaction times under the simultaneous condition, the number of free parameters is one in each response type: either u_s or u_d. These values are proportionality constant. Hence, their different values do not produce different function shapes.

Test implications in the prediction of reaction time. The estimated values of the parameters must meet two test implications as follows: (1) Even when overt "same" reaction time is smaller than "different" reaction time, the estimated value of d_s should be smaller than that of d_d : That is, "same" judgment under the individual point of view (by any individual "describing" procedure) takes a longer time than an individual "different" judgment ($1/a_s > 1/a_d$). To conclude that two descriptions are the same, the individual procedures must confirm that all description units are the same. On the other hand, they can conclude they are "different" when any part of the descriptions mismatch. Each "same" judgment should thus be slower. In the present theory, the "fast-same" phenomenon is not produced by some specific parameter settings, but it is produced by the decision rules across the seven parallel procedures.(2) The estimated value of u_s is greater than that of u_d , since the priming effect exerted by "same" stimulus is greater than that exerted by "different" stimulus.

Predictions of reaction time. Proctor and Hurst(1982: Same-case Blocked condition) and Krueger(1984: Experiment 1) used both the simultaneous- and successive- presentation methods. In addition, they employed a set of stimulus letter pairs which lacked a large heterogeneity of stimulus difference.

Krueger(1984) employed "different" string pairs in which the number of differing letter positions was one. On the one hand, Proctor and Hurst(1982) also used "all-different" pairs, but these were excluded from the present prediction. The reason for this exclusion is that parameter estimation based on only one observed value can not have any empirical significance. In short, predictions based on such an estimation must necessarily fit observed values. Furthermore, though Krueger(1984) used letter string lengths of more than 4, these data were also excluded. The data for these long strings will be mentioned in the next paper, because the present theory assumes that another subprocess (Level 4) describes and tests such long strings in a letter-by-letter manner.

A computer[2] sought the values which minimize the quantity

$$\Sigma\ [RT(L) - E[RT(L)]\ ]^2,$$

where $RT(L)$ represents the observed mean reaction time, and E equals the expected value. The predicted and observed values are shown in Figure 2 and Figure 3. The observed data in the former figure are Proctor and Hurst's(1982) results, and that in the latter are Krueger's(1984).

Table 3 shows the parameter values estimated from Proctor and Hurst's(1982) data. The values of d_s and b were estimated from "same" reaction times under the simultaneous condition, while that of d_d was estimated from "different" reaction times under the same condition. These values (d_s , d_d , and b) were also used to predict reaction times under the successive condition. Thereby, except for the simultaneous-"same" function, the number of estimated values was one in each prediction of individual functions. These values clearly meet both of the two test implications (see Table 3)[3].

Table 4 shows the parameter values employed in the prediction of Krueger's(1984) data. These values were estimated by the method mentioned above, although since the first estimated base time was below 100 msec, a value of 100 msec was used (See Luce, 1986, p.59-64. His Table 2.1 suggests that the value is not below a value of 100 msec: See McCgill, 1963; McClelland, 1979; they adopted the value of 100 msec).

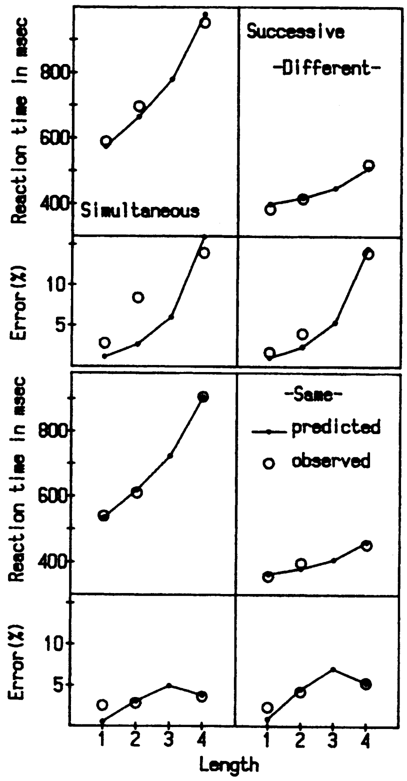

Figure 1. Predicted and observed values: the predicted reaction times are based on Equations 12 and 13 (simultaneous), and on Equations 19 and 20 (successive); the predicted error rates are based on Equations 26 and 28. The observed values are from Proctor and Hurst (1982: Same-case Blocked condition).

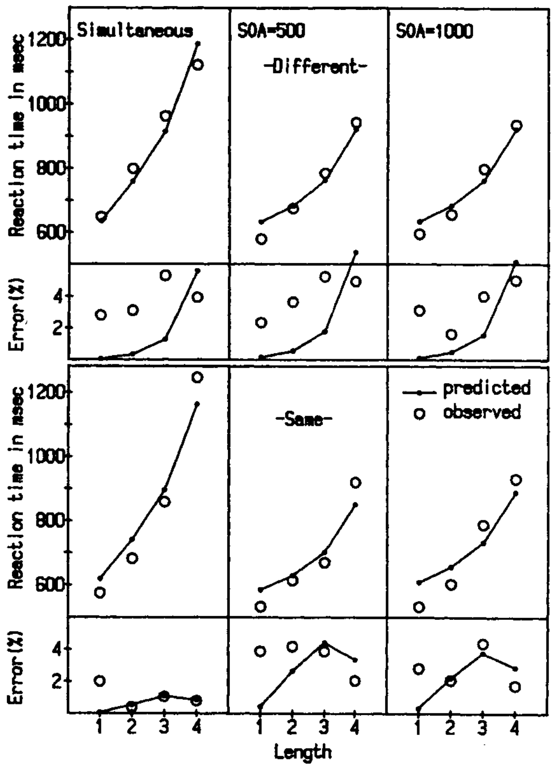

Figure 2. Predicted and observed values: the predicted reaction times are based on Equations 12 and 13 (simultaneous), and on Equations 19 and 20 (SOA=500 and SOA=1000); the predicted error rates are based on Equations 26 and 28. The observed values are from Krueger (1984: Length of less than five in Experiment 1).

Table 1
Parameter values to predict Proctor & Hurst's (1982: Same-case Blocked condition) RT data.

	estimation+	value
Common parameters		
b: base time	simultaneous .same.	164
d_s: discriminability constant for .same. pair	simultaneous .same.	.0004445
d_d: disciminability constant for .diff. pair	simultaneous .diff.	.0073363
Successive condition		
u_s: priming parameter for .same. pair	successive .same.	1.5214025
u_d: priming parameter for .diff. pair	successive .diff.	1.4583898

+ "Estimation" refers to data which was used to estimate each parameter value.

Table 2
Parameter values to predict Krueger's (1984: Length of less than five in Experiment 1) RT data.

	estimation+	value
Common parameters		
b: base time	theoretical	100
d_s: discriminability constant for .same. pair	SOA=0; .same.	.0003039
d_d: disciminability constant for .diff. pair	SOA=0; .diff.	.0054044
SOA=500		
u_s: priming parameter for .same. pair	SOA=500; .same.	.8250413
u_d: priming parameter for .diff. pair	SOA=500; .diff.	.7680102
SOA=1000		
u_s	SOA=1000; .same.	.7828940
u_d	SOA=1000; .same.	.7672302

+ "Estimation" refers to data which was used to estimate each parameter value.

Although the reaction times in Krueger(1984) are generally greater than that in Proctor and Hurst(1982), the base time estimated by the former was smaller than of the latter. The smaller base time was probably caused by a difference in their apparatuses. The value of base time estimated from the data at lengths of 1, 2, and 3 was also somewhat smaller than the value of 100 msec; see below.

Krueger(1984) also employed string lengths of more than 4; this may have caused the greater reaction time. In fact, "same" reaction time for 4-letter strings under the simultaneous condition was greater than "different" one (the disparity is about 120 msec; this seems too large). In contrast, Proctor and Hurst(1982) found a clear "fast-same" disparity for 4-letter strings. This contradiction between these results can be interpreted as follows: The present theory assumes that if letters within a string are randomly selected, the strings longer than 4 are processed by Level 4 in a letter-by-letter manner. Since Krueger(1984) used various string lengths, the subjects became careful when selecting a resolution level, and the mean reaction times for the length of 4 were the weighted sum of the Level 3 and Level 4 latencies. The fits of Krueger's(1984) data are not so good as that of Proctor and Hurst(1982); this could be caused by the mixing of Level 3 and Level 4 latencies. In addition, this mixing could cause the too-small value for estimated base time. (The plots of error rates in the former are rather scattered. This also may be another cause). The parameter values clearly meet both test implications (see Table 4).

It should be noted that these good correlation between predictions and the reaction times confirm the assumption concerning the limit of describable string length: four letters. The definition for the compactness function, *C*, includes this assumption. The value of this function is zero when *L* is 5. Therefore Level 3 is unable to process five-letter strings; its termination latency becomes infinite. Although this is a rather stern constraint, the correlation to data was good. Besides, because of the very small number of free parameters, these strongly support the present model.

Predictions of error rates. In order to calculate false "different"-response probability, we must first consider a probability that one letter within a string is spuriously judged as "mismatch." Let *P* represent the probability of spurious one-letter mismatch, when the "one-letter mismatch" designates that a procedure judged one letter within a string is different from a corresponding-position letter within another string. This probability decreases with longer string length. Even when letters within a string are randomly selected, the string can be described efficiently. Thus, the amount of information about one letter decreases with string length. That is, when a description has redundancy, a more efficient description is possible than the redundant description. Longer strings increase the possibility that one-letter-by-one-letter descriptions have redundancy. For example, a string such as 'EK' could be described as "both have a left-vertical line...." Such a description is simpler than the combination of two descriptions, each of which denotes each letter. Because it is assumed that strings are described as efficiently as possible, the amount of information denoting one letter becomes smaller when string length increases. Then, *P* represents the monotonic decreasing function of *L* (the smaller the number of the parts of a machine is, the smaller the break-down probability of that machine becomes). The present model assumes that the compactness function, *C*(*L*) (Equation 7), represents this decreasing rate (cf., Krueger,1978, adopted Weber-Fechner's law). Accordingly,

$$P(L) = p\, C(L) \qquad L > 0 \qquad (24)$$

when *p* represents the probability of false "different" judgment on an individual description pair in one-letter comparisons ($L = 1$). It should be noted that when *L* equals 0, this equation does not define error probability, since *p* relates to one-letter comparison. Error probability for blank stimuli should be calculated using some other equation which does not involve *p* .

The probability of false "different" judgment for each description pair is

$$1 - \{1 - P(L)\}^L, \qquad (25)$$

since even when any bit of description is changed by internal noise, the description pair is judged as "different" one.

When all of the seven description pairs produced at a trial are different, the categorizer judges that the stimulus pair is "different." In such a case, the probability of false "different" response (error rate at "same" trial: Er_s) is

$$Er_s(L) = [1 - \{1 - P(L)\}^L]^7$$
$$= [1 - \{1 - p\,C(L)\}^L]^7, \qquad (26)$$

Errors during "different" trials are false "same" errors. In order to calculate false "same"-response probability, we must consider first a probability that one letter within a string is spuriously judged as "match." Let Q represent the probability of spurious one-letter match, where the "one-letter match" denotes that a procedure judges one letter within a letter string is different from one corresponding-position letter within another string. This probability increases with string length, because the amount of information about one letter decreases with the length, as mentioned above. That is, a spurious match requires that all mismatching units (symbols) in the description part are changed to identical by internal noise. Hence, the smaller the number of units is, the greater spurious match probability becomes. Then, Q becomes a monotonic increasing function of L. Since this result also depends on compactness, it is assumed that

$$Q(L) = q^{C(L)} \qquad L > 0, \qquad (27)$$

when q represents the probability of false "same" judgment on an individual description pair in one-letter comparisons. It should be noted that q relates to one-letter comparisons, and that when L equals 0, this equation does not define error probability. Since $C(L)$ denotes a monotonic decreasing function less than or equal to 1 ($L > 0$), Q represents a monotonic increasing function of L.

Because "same" response is triggered when any one "same" description pair is found, the probability of false "same" response (error rate at "different" trial: Er_d) is

$$Er_d(L) = 1 - [1 - \{1 - P(L)\}^{L-1}\ Q(L)]^7$$
$$= 1 - [1 - \{1 - p\,C(L)\}^{L-1}\ q^{C(L)}]^7, \qquad (28)$$

This determines that the number of differing letter positions is one. That is, the probability of false "same" judgment on an individual description pair is the product of the following two probabilities: the spurious one-particular-letter match probability and the probability that all L -1 letters are correctly judged to be the same. It should be noted that the "sameness" judgment criterion requires that two descriptions are exactly identical. In short, any small description error produces "different" messages.

The test implication of the parameter values (p and q) is that the estimated value of p should be greater than that of q, even when the observed false "different" rates are less than false "same" ones. That is, Krueger's(1978) theorem states that spurious mismatch probability is greater than spurious match.

Table 3
Estimated parameter values to predict error rates

	Proctor & Hurst (1982)	Krueger (1984)
	simultaneous	
p: false mismatch constant	.47005	.34713
q: false match constant	.00166	.00010
	successive	*SOA=500*
p	.50679	.45843
q	.00154	.00023
		SOA=1000
p		.44383
q		.00017

Values for p were estimated from false "different" rates under individual presentation conditions. These values were substituted for p in Equation 28. Thereby, there was one free parameter estimation in each prediction of individual error rate functions. Table 5 shows the estimated values. It is clear that all estimated value pairs (p and q) under given individual conditions match the test implication.

The estimated values of p are rather large, but this fact is reasonable. That is, even when error probability in each information unit is fairly small, the error probability in an information complex is not small:

Let v represent the error probability of a one-letter description. Then,

$$p = 2\,v(1 - v) + (v^2 - w) \tag{29}$$

where w represents the probability that internal noise causes both descriptions to be changed but to become the same in the end result. Since w is very small[4],

$$p \fallingdotseq 2\,v(1 - v) + v^2 \tag{30}$$

When p equals .50, v is .29. Krueger(1978) stated that the number of features per character is 100: If this is correct,

$$v = 1 - (1 - v_f)^{100} \tag{31}$$

when v_f represents the probability of feature error; then

$$v_f \fallingdotseq .0034$$

The values of v_f which were estimated by Krueger's (1978) model were about .06. The present estimated value of v_f is far smaller than that in Krueger(1978). (Even if the number of features per character is only 20, the value of v_f is .017.) Of course, this discussion is based on the assumption that element (or feature) errors are independent of each other. However, even if it is not true, the value of v_f should not be large.

The fact that a small degree of noise produces a large probability of overall error is the theoretical basis for the exhaustive search assumption of "different" judgment. It should be noted that in recognition tasks, this error (v) will not produce a large overt-error rate: When the categorizer decides a letter name, a completely-correct perceptual description is not needed. In short, it is sufficient to consider main stimulus-letter structures.

Figure 2 shows error rates observed by Proctor and Hurst(1982) and predicted values. The values in Figure 3 are Krueger's(1984) data and predictions. The values shown in the row of "same" trial are the rates of false "different" responses, and that in the row of "different" one are false "same" rates.

There is a dip in the predicted false "different" response ("same"-trial error; Equation 26) at a length of 4. Unfortunately, Proctor and Hurst(1982) did not employed a string length of 3, and the Krueger's(1984) plots of error rate are somewhat scattered. However, it seems that visual inspection reveals the tendency for the observed dip in Figure 2.

It seems that the present decision-rule hypothesis works, though the mathematical model is no more than a first approximation (predicted standard deviations of reaction times are too large, as mentioned above). Applications of the present theory to other matching phenomena and more cognitive tasks will be presented in the next paper; general discussions and a conclusion will also.

Abstract

Saito(1996) proposed a theory of "same"-"different" judgment mechanism, theory which addresses a wide range of problems. In the present paper, a mathematical model describing a subsystem of the theory is constructed. The model approximates a multiple-description (multiple-viewpoint description) system and a rule of search for a correspondence between descriptions. The self-terminating correspondence-search hypothesis explains the "fast-

same" phenomenon. Predictions derived from the mathematical model are fitted to reaction times and error rates in a multiletter matching task. In each prediction of individual reaction-time and error rate functions, the number of free parameters other than base time is one: namely the model has an assumption which states that reaction times and error rates in a pattern matching task do not depend on quantitative judgment-criteria settings but on decision rules across description viewpoints.

References

Baddeley, A.D. & Hitch, G. (1974) Working memory. In G. H. Bower (Ed), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-90). New York: Academic Press.

Eriksen, C.W. & Schultz, D.W. (1978) Temporal factors in visual information processing: A tutorial review. In J.Requin(Ed), *Attention and Performance VII* (pp.3-20). Hillsdale, New Jersey: Erlbaum.

Krueger, L.E. (1978). A theory of perceptual matching. *Psychological Review*, **85**, 278-304.

Krueger, L.E. (1984). Self termination in Same-different judgments: Multiletter comparison with simultaneous and sequential presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 271-284.

Krueger, L.E. (1986). Positive effect of heterogeneity of difference on the same-different disparity in letter matching. *Perception & Psychophysics*, **39**, 117-122.

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

McClelland, J.L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, **86**, 287-321.

McGill, W.J. (1963) Stochastic latency mechanisms. In R. D. Luce, R.R.,Bush, & E.Galanter (Eds), *Handbook of mathematical psychology* (Vol.1: pp.309-360). New York: Wiley & Sons.

Miller, G.A. (1956) The magical number seven, plus or minus two: Some limits of on our capacity for processing information. *Psychological Review*, **63**, 81-97.

Proctor, R.W. (1986). Response bias, criteria settings, and the fast-same phenomenon: A reply to Ratcliff. *Psychological Review*, **93**, 473-477.

Proctor, R.W., & Hurst, P.W. (1982). A parametric investigation of multiletter matches. *Perception & Psychophysics*, **32**, 75- 84.

Proctor,R.W., & Rao,K.V. (1982). Null effects of exposure duration and heterogeneity of difference on the same-different disparity in letter matching. *Perception & Psychophysics*, **33**, 163-171.

Saito,M. (1996) Components in "same"-"different" judgments as an interface between perception and higher cognition (1). *Bulletin of Mimasaka Women's College and Mimasaka Junior College*, **41**, 22-37.

Stevens, S.S. (1957) On the psychophysical law. *Psychological Review*, **64**, 153-181.

Ueno, T. (1976) Temporal summation and reaction time to double-light pulses at suprathrethold levels. *Perception & Psychophysics*, **19**, 399-404.

Footnotes

1. I thank an annoymous reviewer of *Psychological Bulletin*; the reviewer pointed out the possobility of the too-large standard deviation, and I had not carried on any computation of the values.

2. The calculations were done with floating point number of 16 decimal digits.

3. Reaction times of error response is not modeled here; the number of error responses is usually small, and some of them may be an expression of lower vigilance or so forth.

However, a qualitative prediction of false "same" reaction time is easy if all error reaction times depends only on description error. Mean false "same" reaction time will be greater than correct "same" reaction time but smaller than correct "different" one:

During "different" trials, all seven pairs of the descriptions might become the same by chance. Mean false "same" latency in this case is identical to mean correct "same" latency. In another case, one of the pairs might becomes the same by chance, and this spurious match pair may be any one of the pairs; that is, such a spurious match latency may be maximum, minimum, or intermediate at the trial. Consequently, mean latency in this case is greater than that of correct "same" latency, because correct "same" one is an expression of the minimums at individual trials. Besides, this mean is smaller than mean correct "different" latency, because "different" latencies are an expression of the maximums at individual trials. In other cases (two of the pairs become the "same" by chance, or so forth), mean latencies is intermediate: smaller than mean in one-pair case and greater than that in all-pair case.

Overt mean false "same" reaction time is a weighted sum of the latencies in these cases (the all-pair case is rare, and the one-pair case is the most frequent); consequently, it is somewhat larger than mean correct "same" reaction time and fairly smaller than mean correct "different" one.

4. The value of w should not be larger than that of q; when p equals .51, q is .0015 (see Table 5). Even if we assume that w is equal to q , estimated value of v_f is practically identical to the value in the text; the difference between these two values is within a round-off error range.

Appendix A

The probability distribution function of the latency that an individual-viewpoint processing terminates is

$$F(t) = \int_0^t \int_0^{t-y} r(x)\ q(y)\ dx\ dy$$

The integrated form of this equation is obtained on referring to Equations (3) and (4) in the text, as follows:

$$\begin{aligned} F(t) &= \int_0^t \int_0^{t-y} 2(1 - e^{-kx})\, ke^{-kx}\ a\ e^{-ay}\ dx\ dy \\ &= \int_0^t \{1 - e^{-k(t-y)}\}^2\, ae^{-ay}\ dy \\ &= 1 - \frac{2a}{a-k}\, e^{-kt} + \frac{a}{a-2k}\, e^{-2kt} \\ &\quad - \frac{2k^2}{(a-k)\ (a-2k)}\, e^{-at} \qquad \text{(A1)} \end{aligned}$$

Appendix B

The expected value of overt "same" reaction time is

$$E[RT_s] = \int_0^\infty t\, dG(t)\,/\,dt\ \ dt + b\,,$$

The following equation is obtained on referring to Equation (6) in the text:

$$E[RT_s] = \int_0^{\infty} 7\, t\, \{1 - F(t)\}^6\ dF(t)\,/\,dt\ \ dt\ + b$$

$$= [-t\, \{1 - F(t)\}^7]_0^{\infty} + \int_0^{\infty} \{1 - F(t)\}^7\ dt + b$$

In this equation,

$$[-t\, \{1 - F(t)\}^7]_0^{\infty} = 0,$$

because $t\ \{1\text{-}F\ (t)\}^7$ can be rewritten on referring Equation A1 in Appendix A as follows:

$$t \sum_{h+i+j=7} \frac{7!}{h!\ i!\ j!}\ A^h\ B^i\ C^j$$

$$\times\ \exp\{-\,(hk + 2ik + ja)\ t\}$$

where $A = 2a\,/\,(a - k)$; $B = -a\,/\,(a - 2k)$; and $C = 2k^2\,/\,(a$ - k) $(a - 2k)$. Every term of this polynomial expression is a product of t, a constant, and an exponential function; the last is $\exp\{-\,(hk + 2ik + ja)\ t\}$.
Then if $\lim_{t\to\infty} t\ \exp(-xt) = 0$ (x is an arbitrary positive number), $[-t\{1\text{-}F(t)\}^7]_0^{\infty} = 0$: Since $\lim_{t\to\infty} t = \infty$, and $\lim_{t\to\infty} e^{xt} = \infty$ also,

$$\lim_{t\to\infty} \frac{t}{e^{xt}} = \lim_{t\to\infty} \frac{dt\,/\,dt}{de^{xt}\,/\,dt} = 0.$$

Then, the integrated form of the expectancy value is obtained as follows:

$$E[RT_s] = \int_0^{\infty} \{1 - F(t)\}^7\ dt\ + b$$

$$= \int_0^{\infty} \sum_{h+i+j=7} \frac{7!}{h!\ i!\ j!}\ \left(\frac{2a}{a-k}\right)^h$$

$$\times \left(-\frac{a}{a-2k}\right)^i \left(\frac{2k^2}{(a-k)\ (a-2k)}\right)^j$$

$$\times\ e^{-(hk+2ik+ja)t}\ dt + b$$

$$= \sum_{h+i+j=7} \frac{7!}{h!\ i!\ j!}\ \left(\frac{2a}{a-k}\right)^h$$

$$\times \left(-\frac{a}{a-2k}\right)^i \left(\frac{2k^2}{(a-k)\ (a-2k)}\right)^j$$

$$\times \frac{1}{hk+2ik+ja} + b\,.$$

Appendix C

The expected value of overt "different" reaction time is

$$E[RT_d] = \int_0^{\infty} t\ dH(t)\,/\,dt\ \ dt + b$$

$$= \int_0^{\infty} 7\, t\, \{F(t)\}^6\ dF(t)\,/\,dt\ \ dt + b$$

$$= -\sum_{\substack{g+h+i+j=7 \\ (g \neq 7)}} \frac{7!}{g!\ h!\ i!\ j!}\ \left(-\frac{2a}{a-k}\right)^h$$

$$\times \left(\frac{a}{a-2k}\right)^i \left(-\frac{2k^2}{(a-k)\ (a-2k)}\right)^j$$

$$\times \frac{1}{hk+2ik+ja} + b\,.$$

This integrated form is obtained with the idea mentioned in Appendix B. Note that a term (g=7) is excluded. When g equals 7, the term is a constant of 1; then this becomes t by the integration, and $[t\ F^7(t) - t\ ]_0^{\infty} = [-t\ \{1\text{-}F^7(t)\}]_0^{\infty} = 0$, because of the same reason mentioned in Appendix B.

Appendix D

$$F^*(t) = \int_0^t \int_0^{t-y} r^*(x)\ q(y)\ dx\ dy$$

$$= \int_0^t \int_0^{t-y} k\ e^{-kx}\ a\ e^{-ay}\ dx\ dy$$

$$= 1 - \frac{k}{k-a}\ e^{-at}\ + \frac{a}{k-a}\ e^{-kt}.$$

(1996年12月2日　受理)