

The effect of uncertainty in stimulus perception on same-different judgments

DAVID NAVON

University of Haifa, Haifa, Israel

Existing models of same-different judgments are examined as to how useful they are in situations in which encoding of stimuli is presumably imperfect. Subjects were asked to judge whether a probe stimulus composed of one or four figures was identical to the corresponding components of a target stimulus that was always composed of four figures. It was found that correct "different" responses were faster than correct "same" responses, that correct "same" responses were faster than either incorrect "same" responses or incorrect "different" responses, and that reaction time to probes of four figures was shorter than reaction time to probes of one figure. These results and others are interpreted as embarrassing for existing models of matching behavior.

Numerous models have been suggested to account for behavior in same-different judgment situations with sequential presentation. They are reviewed and discussed in various places (e.g., Audley, 1973; Bamber & Paine, 1973; Nickerson, 1972). Most of them focus mainly on the processes of retrieval, comparison, and decision involved in the binary choice that the subject has to make. Little attention has been given to the effect encoding of the target stimuli might have on the stage of decision. The results reported in this paper demonstrate that this neglect can be detrimental when judged stimuli are not so readily encoded.

Existing models for same-different judgments can be roughly classified in two categories that may be called analytic models and holistic models.

Analytic models share at least four assumptions: (1) Decomposability—A visual percept comprises of a finite number of elements. (2) Quantal logic—Each of the elements is either perceived or not, all or none, and, if perceived, either retained or not by the time of the test (in a typical same-different judgment situation, it is assumed that all elements are both perceived and retained). (3) Passive memory—During the period between the presentation of the target and the probe, there is no distortion of the memory traces of the elements and there is no intrusion of unperceived elements into the traces; thus, the only type of memory deterioration possible is loss of elements. (4) Element matching—Comparison is done by matching corresponding elements in the target and the probe, so that one or more mismatches results in a response of "different."

The specific models within this category may differ with respect to at least six properties: (1) the assumed complexity of the elements (e.g., line segments, letters), (2) the particular mechanism hypothesized to perform the comparison, (3) whether the comparison process is parallel or serial, (4) whether all retained relevant

elements are compared or just a subset of them, (5) whether comparison within that subset (or the whole set) is exhaustive or self-terminating, and (6) whether decision is biased in favor of "same" responses or in favor of "different" responses, or is unbiased.

The second category of models, holistic models, postulates the existence of a holistic template match process that may operate alone (Sekuler & Abrams, 1968), in parallel with analytic comparison processes (Bamber, 1969), or as a first comparison attempt that is followed by analytic comparison only if the holistic approach fails (Marcel, 1970; Tversky, 1969). Most of the advocates of such models assume that the holistic process can output only "same" responses, probably in order to account for the pervasive and somewhat puzzling finding that latency of "same" responses is usually shorter than latency of "different" responses. However, since the output of matching is presumably some measure of goodness of fit, it is not clear why a holistic process cannot detect difference in case of a very bad fit, as proposed by Atkinson and Juola (1974). If it does, however, then one begins to wonder what is accomplished by assuming a holistic match process, other than accounting for the particular case in which "same" or "similar" stimuli are generally "highly" similar, yet dissimilarity of "different" stimuli is "moderate" on the average, thus has to undergo analytic comparison before decision. That even this prediction may fail is demonstrated later.

In order to examine how robust existing models may be in situations in which stimulus encoding is limited (by familiarity, complexity, or duration of the stimulus, or by subjects' attention), I designed an experimental task with three essential features: (1) The probe is sometimes a whole figure of the type used as target stimuli, and sometimes just part of that figure. (2) The stimuli are unfamiliar and moderately complex patterns.

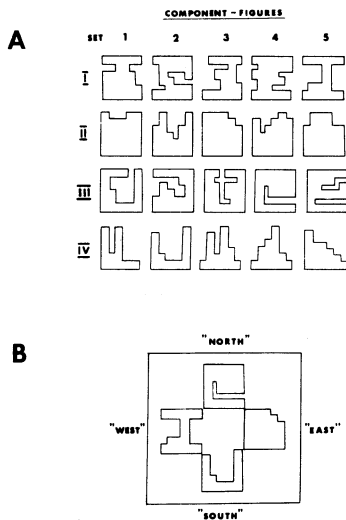


Figure 1. (A) The five sets of component figures used to generate the stimulus patterns. (B) Example of a possible target pattern.

(3) The target stimulus is presented for a brief duration.

METHOD

Stimuli and Design

A target stimulus was composed of four component figures, each taken from a different row in Figure 1A, laid out in a fashion depicted in Figure 1B. The selection of the four component figures that appeared on a given trial was done randomly by a computer, as was the assignment of component figures to positions. The probe stimulus was a whole pattern about half of the time (whole probe) and just one component figure in one of the positions in which it could have appeared within a pattern the rest of the time (part probe). In about half of the part probes, the figure displayed was identical to the figure that appeared in the same position in the target pattern. In the remaining part probes, the figure was one of the other four figures that could have appeared in the same position, so it was different from all the figures that appeared in the target pattern. In about half of the whole probes, the pattern displayed was identical to the target pattern; in the remaining half, it was different. The "different" whole probe contained two, three, or four component figures, which were different from the figures in the target, with respective probabilities of .55, .36, and .09.

The subjects were instructed to respond "same" only to probes all the components of which appeared in the target.

Five payoff matrices were used. The payoff for any incorrect response was always zero. The payoffs for a correct "same" response and a correct "different" response were 2.8 and 1.0 cents, 2.2 and 1.4 cents, 1.8 and 1.8 cents, 1.4 and 2.2 cents, or 1.0 and 2.8 cents. Each payoff matrix was used in three consecutive sessions. The symmetrical payoff matrix was used first for each subject. The order of the other matrices was counterbalanced across the subjects.

Subjects

Four students of the University of California, San Diego, served as subjects and were paid for their services.

Apparatus

The equipment consisted of a display Tektronix oscilloscope and a keyboard, both controlled by a PDP-9 computer in an

acoustically isolated booth. The subject, who was run individually, sat in the booth in front of a table on which the keyboard was positioned. The display oscilloscope was positioned at about eye level. Viewing distance was about 45 cm. The side of a square circumscribing a component figure subtended about 2 deg 26 min visual angle.

Procedure

A trial began by a .5-sec warning beep, followed by a small fixation dot that was visible for 1 sec. Following that, the target pattern was displayed for a duration that had been adjusted before the experiment for each individual subject to a level of 75%-80% correct responses. The durations used were 360, 400, and 500 msec. The target pattern was followed by a 100-msec mask that consisted of a grid that overlaid all the possible line segments in the patterns. The probe stimulus followed immediately after the mask and was terminated by the subject's response. The subject responded by pressing one of two keys to indicate "same" or "different." Accuracy and time from onset of the probe stimulus to response were recorded for each trial. Each subject served in 15 sessions administered on different days, and each of the sessions consisted of 250 trials. No subject had more than one session in 1 day. The subjects were told before the experiment about the probabilities for the various probes, and also that the number of different patterns was larger than they might expect to remember.

RESULTS

All the relevant data are presented in Table 1. The important effects are summarized below. Most of the comparisons were done separately for each of the

Table 1
Mean Reaction Times (MRT) in Milliseconds and Proportions (P) of Correct and Incorrect Responses in "Same" and "Different" Trials

Subject	Response	Probe	"Same" Trials		"Different" Trials		Overall MRT
			P	MRT	P	MRT	
1	Correct	Whole	.942	746	.835	659	702
		Part	.645	902	.766	828	863
	Incorrect	Whole	.058	1672	.165	887	1073
		Part	.355	994	.234	1071	1022
2	Correct	Whole	.895	734	.737	722	728
		Part	.695	974	.661	932	954
	Incorrect	Whole	.105	1070	.263	843	902
		Part	.305	1044	.339	1172	1108
3	Correct	Whole	.965	532	.677	567	547
		Part	.551	719	.738	677	696
	Incorrect	Whole	.035	751	.323	594	608
		Part	.449	756	.262	845	787
4	Correct	Whole	.955	819	.664	742	785
		Part	.577	1674	.717	1580	1624
	Incorrect	Whole	.045	2077	.336	854	985
		Part	.432	1898	.283	2948	2297
Total	Correct	Whole	.939	707	.728	672	690
		Part	.617	1062	.720	1000	1028
	Incorrect	Whole	.061	1354	.272	779	874
		Part	.383	1184	.280	1524	1320

Note—Mean RTs and proportions correct and incorrect are tabulated according to type of probe for each subject separately and across subjects.

subjects by applying *t* tests to the data of each. All the comparisons were significant at the .001 level, unless otherwise stated.

(1) Subjects responded faster to whole probes than to part probes. This was true for both correct responses and incorrect responses, with the exception of the incorrect responses of Subject 1. It was also true when correct and incorrect responses were pooled together.

(2) Correct "different" responses were generally faster than correct "same" responses. This was true for whole-probe trials and for part-probe trials for each of the subjects (except for Subject 3 in whole probe).

(3) Correct "same" responses were faster than incorrect "same" responses (*viz.*, incorrect responses in "different" trials). This was true for whole-probe and part-probe trials, for each of the subjects.

(4) Proportion of incorrect responses in whole-probe "same" trials was above the response-confusion level. The subjects responded incorrectly to between 3.5% and 10.5% of the whole-probe "same" trials. In order to make sure that those errors were not due solely to associating the response and the key in the wrong way, the subjects ran through a posttest of 50 trials, in which the target, as well as the probe, was made of just one component figure. The subjects reported that the errors they made, if any, were response-confusion errors. Error rates were .00, .04, .02, and .00, for Subjects 1-4, respectively; these were considered to be the response-confusion levels. The binomial probability that the observed incorrect "different" responses in the test itself arose from response confusion is smaller than .01, both when tested for each subject separately and across all the subjects.

(5) Correct responses were made faster than incorrect responses on "same" trials. Correct "same" responses were faster than incorrect "different" responses for both whole probes and part probes.

(6) Proportion of errors in part-probe "same" trials was generally higher than proportion of errors in part-probe "different" trials. This was true for Subjects 1, 3, and 4 (with *p* values in a *Z* test for difference between proportions smaller than .001), but false for Subject 2 (for his data, the test was nonsignificant: $Z = -1.38$).

DISCUSSION

Can any of the existing models accommodate these findings? Evidently, most analytic models do not predict the finding of some misses in whole-probe "same" trials (Finding 4). Any information extracted from the target is present in the whole probe; so, if there is no memory distortion, the subject should respond "same" unless he is biased against it. However, there are at least two good reasons to dismiss the bias hypothesis. One, the design was supposed to counterbalance motivational factors by using five different payoff matrices. Two, if the bias is implemented via a rule that with some probability *g* responds "different," despite evidence that the probe might be identical to the target, then there should not be any difference between mean reaction times of incorrect and correct responses in "same" trials, but in fact there was (Finding 5).

Whereas the presence of missed whole-probe "same" trials is inexplicable by analytic models unless a bias assumption is involved, failures to report identity are understandable in part-probe trials, because the visual information sampled during the presentation of the target may not overlap at all with the actual probe stimulus. In this case, the subject is assumed to guess. However, if we assume again that the design of this experiment controlled for any motivation to favor one response over the other, then how can an analytic model account for the considerable bias found to respond "different" (Finding 6)?

While the presence of missed whole-probe "same" trials could be interpreted as indicating bias against "same" responses, the finding that correct "same" responses are faster on the average than incorrect ones (Finding 3) conflicts with any analytic model, except for one that is biased in favor of "same" responses. If a "same" response arises from a situation where identity is found in all the element pairs compared, then, since the subject does not have any way to know what the true state of the world is, there should not be any difference between the latencies of incorrect and correct "same" responses. To account for this finding, one might assume that subjects may emit a "same" response even when aware of a discrepancy between the target and the probe, and that the "same" response is faster in "same" trials because matching evidence is more compelling or accumulates faster in those trials. But why should a subject who trusts his perceptual processes and short-term memory tolerate any mismatch at all? A possible reason might be that the decision process is probably tailored to the output of a perceptual process with considerable uncertainty; it acts on the assumption that this output may not be totally veridical or precise. This, however, is inconsistent with the basic postulates of existing models.

Any analytic model, except for one that assumes parallel element matching with constant and equal durations, is challenged by the finding that whole probes are faster to process than part probes (Finding 1). The expected number of elements to be compared is greater in whole-probe trials. One may argue that mean part-probe reaction time is inflated by part-probe trials in which the probe does not contain any elements that correspond to visual information sampled from the target. However, this contention seems fairly untenable in view of the fact that the finding holds even when incorrect responses are eliminated.

Contrary to a prevalent finding in same-different experiments, it was found here that a correct "different" response was faster than a correct "same" response (Finding 2). This result is consistent with the prediction from most conceivable analytic models, since difference can be detected at any phase of comparison, whereas a certain criterion has to be met in order to decide about "sameness." However, this finding conflicts with what is implied by existing holistic models. One can patch those models post hoc by a restriction on the holistic template match to be able to output only "different" responses, but this seems as arbitrary as the present restriction, which assumed just the opposite. This finding could have been reconciled with the model of Atkinson and Juola (1974), if the stimuli in "different" trials had been typically very dissimilar, whereas stimuli in "same" trials had been just moderately similar, so that dissimilarities could be detected in the holistic stage but similarities had to be left for the later analytic one. However, the stimuli in this experiment appear to exhibit just the opposite properties: "Same" stimuli are identical and "different" stimuli usually resemble each other considerably.

In summary, the set of findings in this experiment poses difficulties for any of the existing models of behavior in same-different tasks. This suggests that one or more of their basic assumptions must be unwarranted, at least for the situation reported. Of the four assumptions underlying analytic models that are listed in the introduction, the first and the last were relaxed by holistic models, but those models do not seem to

be more explanatory than the analytic models. Thus, I suspect that there is reason to doubt the assumptions of quantal logic and passive memory. Perhaps future theorization should not overlook the possibility that some uncertainty in representation of target elements is brought about by processes of perception or memory.

REFERENCES

- ATKINSON, R. C., & JUOLA, J. F. Search and decision in recognition memory. In D. H. Krantz, R. C. Atkinson, D. C. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology*. San Francisco: Freeman, 1974.
- AUDLEY, R. J. Some observations on theories of choice reaction time: Tutorial review. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.
- BAMBER, D. Reaction times and error rates for "same"- "different" judgments of multidimensional stimuli. *Perception & Psychophysics*, 1969, **6**, 169-174.
- BAMBER, D., & PAINE, S. Information retrieval processes in "same"- "different" judgments of letter strings. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.
- MARCEL, A. J. Some constraints on sequential and parallel processing and the limits of attention. In A. F. Sanders (Ed.), *Attention and performance III*. Amsterdam: North Holland, 1970.
- NICKERSON, R. S. Binary-classification reaction time: A review of some studies of human information processing capabilities. *Psychonomic Monograph Supplements*, 1972, **4**, 275-317.
- SEKULER, R. W., & ABRAMS, M. Visual sameness: A choice time analysis of pattern recognition processes. *Journal of Experimental Psychology*, 1968, **77**, 232-238.
- TVERSKY, B. Pictorial and verbal encoding in a short-term memory task. *Perception & Psychophysics*, 1969, **6**, 225-233.

(Received for publication August 7, 1978.)