Experimental test of the relation of confidence with the time errors in the method of constant stimuli

SERGIO CESARE MASIN
University of Udine, Udine, Italy

and

GIANPAOLO CERETTA
University of Padua, Padua, Italy

The experiment reported here tested the common belief that comparative judgments become biased because response habits intervene when the observer feels uncertain about which response to give. Using the method of constant stimuli, observers were shown a train of lines, each separated by a 5-sec interval. The observer compared the length of each line with the length of the line immediately preceding it. In addition to making the comparative judgment, the observer had to rate the degree of confidence he/she had in determining that the last line seen was different in length from the immediately preceding line. The results showed that (1) degree of confidence was progressively lowered (slightly) as the number of responses increased, (2) degree of confidence was higher when the last line seen was a variable (even if the observer did not know which was the standard or the variable), and (3) time errors also occurred when the observer was completely certain that there was a difference in length between the compared lines. It may be concluded that the observer's state of uncertainty did not activate any preexisting response habit. The results seem to fit an inferential-decision theory of comparative judgment.

When the method of constant stimuli (MCS) is used to determine the point of subjective equality (PSE), the observer is shown a standard (S) stimulus paired with one of a number of variable (V) stimuli. The set of S-V pairs presented one at a time in random order is called a *series*. Many successive series are shown, and the observer has to report which of the two stimuli in each pair of a series is greater than the other with respect to a given attribute. The PSE is defined as the value on the stimulus continuum of Vs that has 50% probability of being reported as greater than S.

The difference PSE−S is a bias measure called the *constant error*. In the MCS, S and V may be presented simultaneously or successively. When S and V are successive, the constant error is called the *time error*. A time error may occur when V is compared with the immediately preceding S. When S is compared with the immediately preceding V, another time error may occur. The signs of the two time errors are opposite (Fechner, 1860). Martin and Müller (1899) found that the two time errors may have a different magnitude (see Needham, 1934, for a review of early studies). This result has been confirmed by Woodruff, Jennings, and Rico (1975). The occurrence of the two opposite time errors is referred to as the *presentation-order effect*.

In a number of important studies, the time errors in the MCS have been ascribed to biasing response habits, such as response preferences, the tendency to avoid repeating the same response, and so forth (Guilford, 1954; Hellström, 1985). A biasing habit would be activated when the observer feels uncertain about which response to give. Referring to the MCS, Guilford (1954) wrote that "such habits come into play most when [the observer] feels there is little basis for making a judgment and when he senses the chance atmosphere of the situation" (p. 150). (If the habit were always present, then it should occur also when the difference between the compared stimuli is well above the differential threshold. This presence would imply the counterfactual conclusion that the same number of judgmental errors would be produced when the stimuli are perfectly discriminable.)

To date, Guilford's (1954) assumption has not been tested experimentally, probably because it seems intuitively convincing. However, there is no rationale for it. A test of this assumption seems important. In fact, if experimental results were to indicate that the observer's state of uncertainty is unrelated to the magnitude of the time errors, then time-error theories relying on the concept of a biasing habit would be weakened in that there would be no plausible reason why the biasing habit becomes activated in the observer.

Absence of a relationship between the time errors and the subjective state of uncertainty obviously means that the time errors occur both when the observer is uncer-

tain and when he/she is completely certain that he/she perceived a difference between the compared stimuli. Thus, if the time errors also occur when the observer is completely certain that he/she perceived a difference, then it should be concluded that the time errors must have already occurred before the observer produced the overt response. This possibility would be in favor of a memory theory of the time errors. However, Masin (1987) and Masin, Mazzoni, and Vallortigara (1987) showed experimentally that the time errors cannot be explained in terms of a comparison of the perceived stimulus with a memory trace, or with internal standards such as the adaptation level. Masin, Fanton, and Crestoni (1988) also showed that the asymmetry effect occurs for both the time and space errors, thus strongly indicating that a purely memory explanation of the time errors cannot apply.

Masin and Fanton (1988) proposed a decision theory of the time errors in which they suggested that the comparison judgment of two stimuli with respect to a given attribute derives from the comparison of the categories, on a ''less-greater'' scale, in which the stimuli would be assigned automatically (Bower, 1971). They explained the time errors in terms of an inferential decision process that would sometimes control the overt response on the basis only of the category in which the last stimulus presented is assigned, because the last category receives more weight or because the category relative to the preceding stimulus cannot momentarily be retrieved from memory. This theory implies that the observer's state of uncertainty is a consequence of the categorization of the stimuli. The observer may be certain that two given stimuli were different when the second stimulus is (stochastically) assigned to a very extreme category (independent of the physical or perceptual difference). Accordingly, in this case, no relation between subjective uncertainty and magnitude of the time errors is expected.

The purpose of the experiment reported here was then to achieve a test of Guilford's (1954) assumption.

METHOD

Observers

The observers were 50 university students who were asked to participate in the experiment when they entered the General Psychology Department at the University of Padua.

Stimuli

The stimuli were moderately bright orange horizontal lines displayed for 1.5 sec on a nonluminous dark-gray 17×23 cm monitor screen controlled by an Apple microcomputer. The center of each line coincided with the center of the monitor screen. Each line was signaled 1.5 sec before its appearance by a very brief sound. The viewing distance was 116 cm. The stimulus conditions were substantially a repetition of those used by Masin and Fanton (1988). The MCS was applied by presenting S and V in pairs, S first and V second. There were 10 series. The inter-stimulus and interpair intervals were both set at 5 sec, thus producing a train of lines in which no temporal grouping of lines was possible. The Vs, hereinafter called V1–V5, were five lines having lengths of 87.2, 89.9, 92.6, 95.3, or 98.0 mm on the monitor screen. S was set at 95.3 mm, which matched V4.

To prevent the observer from seeing images reflected off the monitor screen, the experimental room was dimly illuminated (0.6 lx at the monitor screen), and the observer viewed the stimuli through a 9×13 cm
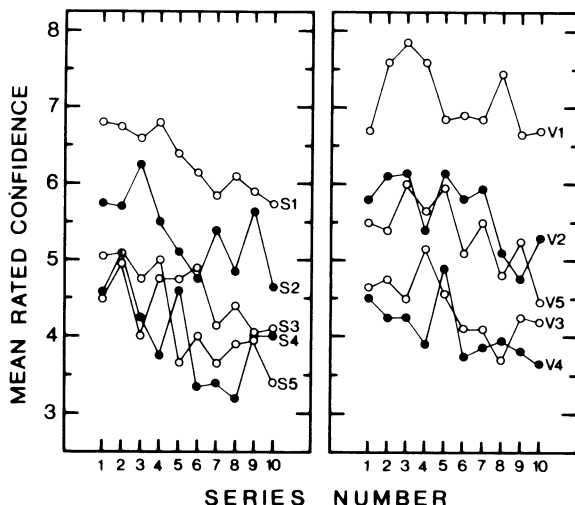

Figure 1. Mean rated confidence, relative to the standards (left) and to the variables (right), as a function of the series number.

hole in a 75×100 cm dark screen placed 63 cm from the monitor screen. The eye level was adjusted through a head- and chinrest. The stimuli were viewed binocularly.

Procedure

The observer's task was to report whether each line on the monitor screen was longer or shorter than the immediately preceding line. ''Equal'' and ''don't know'' responses were not allowed. Immediately after each ''longer'' or ''shorter'' response was given, the observer had to rate how much confidence he/she had in determining that the last stimulus was different in length from the immediately preceding stimulus. The rating of confidence was made using integers in the range from 0 (lack of any confidence) to 10 (absolute certainty that there was a difference in length). Observers started responding after the second line appeared. To compensate for the absence of any response to the first line (an S), a final S was presented at the end of the 10 series. Thus, each observer judged 100 stimuli, 50 Ss and 50 Vs. Observers were not told which were Ss and which were Vs.
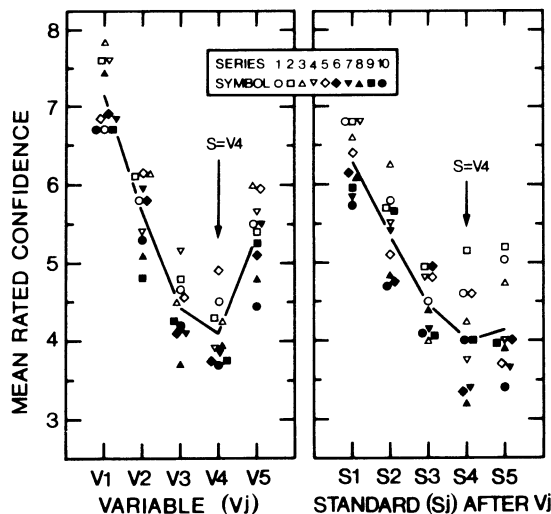

Figure 2. Mean rated confidence as a function of the position of the variable in the ordered series of variables (left), and of the position of the standard immediately following the variable in the ordered series (right).

RESULTS AND DISCUSSION

The results from the confidence rating are depicted in Figures 1 and 2. The mean rated confidence is represented as a function of the series number in Figure 1, relative to the responses given to the Ss (left) and Vs (right). Figure 2 shows how the mean rated confidence varies as a function of the position in the ordered series of Vs (left) and Ss (right). In both figures, the symbols S1, S2, ..., S5 represent the standards immediately following the variables V1, V2, ..., V5, respectively. A 2 (kind of stimulus: S vs. V) × 5 (position in the ordered series) × 10 (series number) analysis of variance (ANOVA) showed that the main effects due to the factors kind of stimulus and series number were statistically significant [$F(1,49)$ = 24.89, $p < .00005$, and $F(9,441)$ = 5.75, $p < .00005$, respectively]. The interaction of kind of stimulus and position in the ordered series was significant [$F(4,196)$ = 8.29, $p < .00005$], thus showing that the two graphs in Figure 2 are statistically nonparallel. The interactions between series number and position in the ordered series, between kind of stimulus and series number, and between all three factors were nonsignificant [$F(36,1764)$ = 1.09, $F(9,441)$ = 1.56, and $F(36,1764)$ = 1.20, respectively].

Thus, based on the ANOVA, it may be concluded that the difference judgments have a mean rated degree of confidence that decreases slowly as the series number increases (perhaps because the observer tends to deploy progressively less attention to the stimuli). It may also be concluded that the observer is slightly more confident when he/she judges a V than an S. Therefore, something peculiar to S, or V, occurs when the judgment is made,

Figure 4. Standard deviation of the psychometric functions represented in Figure 3 as a function of the degree of confidence, for the responses given to the standards and to the variables.
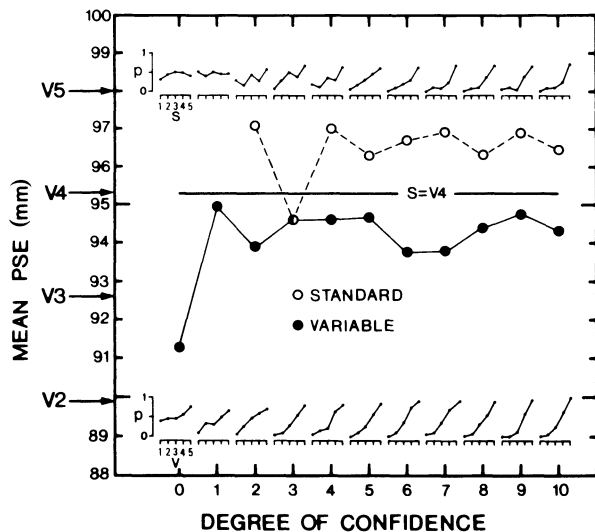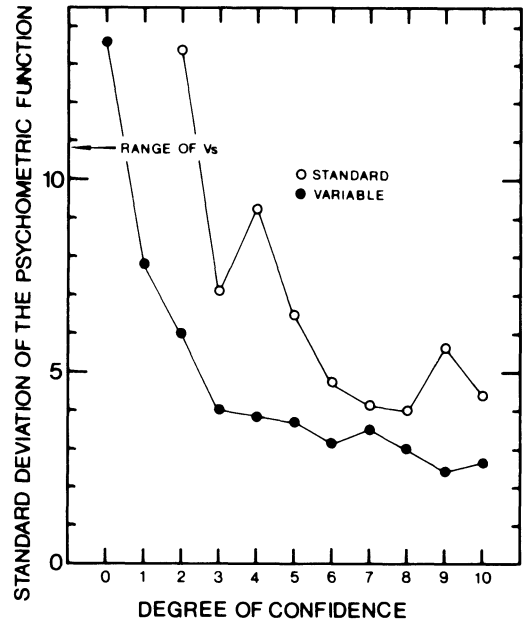
Figure 3. Mean point of subjective equality (PSE), obtained from the responses to the standard (upper graph) and to the variable (lower graph), as a function of the degree of confidence. The psychometric functions for each of the degrees of confidence also appear in this figure. The important result is that the presentation-order effect occurred independent of the degree of confidence.

even if the observer does not know which is S or V. When a comparative-judgment task is executed, the compared stimuli would be automatically classified in categories on a short-long scale (Bower, 1971; Masin & Fanton, 1988). It seems reasonable that the more frequently a category is used, the more hesitant the observer is in categorizing those stimuli that are objectively more likely to be assigned in that category. Ss are the most frequent stimuli, and accordingly result in a less confident judgment.

The constant error computed using the mean PSE, determined by fitting (using the weighted least-squares method) the psychometric function from the "longer" responses given to the Vs, was −1.12. The responses alternative to a response given to an S may be considered equivalent to an implicit response given to the immediately preceding V. (For example, a "shorter" response to S is equivalent to an implicit "longer" response to the preceding V.) The constant error from the implicit "longer" responses to the Vs was 1.20. Both constant errors were statistically different from zero ($z = -5.2$, $p < .00005$, and $z = 2.7$, $p < .005$, respectively). This result confirms the existence of the presentation-order effect when the observer does not know which is S or V, and when the S-V pairs cannot be distinguished because they are phenomenally eliminated in a train of stimuli (Masin & Fanton, 1988).

The "longer" responses to Vs, and the "shorter" responses to Ss (from which the implicit "longer" responses to Vs were derived), were selected for each of the 11 degrees of confidence rating. The mean PSE was computed from the responses so selected by fitting each psychometric function using the weighted least-squares

method. Figure 3 shows the relation between the mean PSE and the degree of confidence. The filled (empty) dots represent the results from the explicit (implicit) "longer" responses to Vs. The psychometric functions relative to the explicit (implicit) responses to Vs are depicted in the lower (upper) part of the figure for each degree of confidence. As may be seen, the psychometric functions relative to S (implicit responses to V) are flatter than those relative to the Vs. (The first two points of the graph relative to S are omitted, given the nearly complete randomness of the implicit "longer" responses for the degrees of confidence 0 and 1.) Figure 4 shows how the standard deviation of the psychometric functions that are represented in Figure 3 varies as a function of the degree of confidence.

The most important result (evident in Figure 3) is that the mean PSE remained substantially unchanged as the degree of confidence varied. The presentation-order effect occurred even when the observer was absolutely certain that he/she perceived a difference in length between V and the preceding S, or between S and the preceding V. Therefore, it may be concluded that the time errors already occurred when the observer produced the overt response. Thus, this result weakens response-bias theories of the time errors, whereas it corroborates Masin and Fanton's (1988) inferential-decision theory.

REFERENCES

BOWER, G. (1971). Adaptation-level coding of stimuli and serial position effects. In M. H. Appley (Ed.), *Adaptation-level theory* (pp. 175-201). New York: Academic Press.

FECHNER, G. T. (1960). *Elemente der Psychophysik*. Leipzig, East Germany: Breitkopf & Hartel.

GUILFORD, J. P. (1954). *Psychometric methods*. New York: McGraw-Hill.

HELLSTRÖM, A. (1985). The time-order error and its relatives: Mirros of cognitive processes in comparing. *Psychological Bulletin, 97*, 35-61.

MARTIN, L. J., & MÜLLER, G. E. (1899). *Zur Analyse der Unterschieds-emfindlichkeit*. Leipzig, East Germany: Barth.

MASIN, S. C. (1987). Different biases in the methods of constant and single stimuli. *Bulletin of the Psychonomic Society, 25*, 379-382.

MASIN, S. C., & FANTON, V. (1988). *An explanation for the presentation-order effect in the method of constant stimuli*. Manuscript submitted for publication.

MASIN, S. C., FANTON, V., & CRESTONI, L. (1988). *An experimental study of the asymmetry effect in the method of constant stimuli*. Manuscript submitted for publication.

MASIN, S. C., MAZZONI, G., & VALLORTIGARA, G. (1987). The first five responses in the method of constant stimuli. *Canadian Journal of Psychology, 41*, 80-83.

NEEDHAM, J. G. (1934). The time-error in comparison judgments. *Psychological Bulletin, 31*, 229-243.

WOODRUFF, B., JENNINGS, D. L., & RICO, N. L. (1975). Time error in lifted weights as affected by presentation order and judgment mode. *Perception & Psychophysics, 18*, 98-104.