Confidence in stimulus predictions and choice reaction time*

E. SCOTT GELLER† and CHARLES P. WHITMAN

Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

In two choice reaction time (RT) experiments, a stimulus prediction and a confidence judgment in the prediction preceded each occurrence of one of two stimulus alternatives. Ss identified each stimulus presentation by pressing a left-hand or right-hand telegraph key. In Experiment I the source of the stimulus predictions and confidence estimates was varied between groups of 20 Ss. For each condition, RT to correctly predicted stimuli was an inverse function of prediction confidence. Following incorrectly predicted stimuli, RT was not reliably influenced by confidence when S gave both predictions and confidence judgments; but RT to incorrectly predicted stimuli was an increasing function of confidence when E verbalized the predictions and confidence estimates or when S predicted and E indicated confidence. In Experiment II Ss made predictions and the validity of E's confidence estimate was manipulated between Ss. When E's confidence was perfectly related to the probability of a correct prediction, choice RT to nonpredicted stimuli was inversely related to confidence. However, choice RT to nonpredicted stimuli was not affected by prediction confidence when E's judgments were random.

Several researchers have concluded that the time necessary to make a correct identification response in a two-choice reaction time (RT) paradigm with one to one mapping of stimuli and responses is dependent on Ss' expectancy for the stimulus presentation. For example, if one stimulus alternative is presented more often than the other, Ss should expect that stimulus more often, and indeed RT is shorter to the more probable stimulus (e.g., Hinrichs & Craft, 1971; Laming, 1969; Remington, 1969). When required to predict verbally each sequential presentation in a two-choice RT task, Ss react markedly faster to correctly predicted (i.e., expected) events than to incorrectly predicted events (e.g., Bernstein & Reese, 1965; Hale, 1967; Keele, 1969; Williams, 1966). These separate demonstrations of probability and prediction outcome effects on two-choice RT may be accounted for with an expectancy model that assumes Ss' readiness to be a dichotomous process, i.e., Ss are either prepared or unprepared for a particular event (cf. Falmagne, 1965; Theios & Smith, 1972).

Recent analyses of stimulus probability and prediction outcome effects on choice RT have supported a notion that Ss are perhaps not only set or unset for a given stimulus but also that Ss' expectancy for a stimulus alternative may vary in degree within a set or unset state. For example, a continuous effect of probability on two-choice RT may have occurred when the probability effect (i.e., faster responses to the more probable stimulus) was an increasing function of the extent of the frequency imbalance between the stimuli of different sequences (Hinrichs & Craft, 1971; Kanarick, 1966; Laming, 1969) or when RT fluctuated

intrasequentially according to trial by trial variations of event probability (DeKlerk & Oppe, 1970; Geller & Pitz, 1970). Moreover, when choice RTs were classified according to both relative frequency and prediction outcome, RTs were faster following correct than incorrect predictions, and, within each prediction outcome category, RTs were faster to the more probable stimulus (Geller, Whitman, Wrenn, & Shipley, 1971; Geller, Whitman, & Farris, 1972; Hinrichs & Craft, 1971). Geller et al hypothesized that Ss were most prepared after predicting the more frequently occurring stimulus and consequently reacted fastest when correctly predicting the more probable stimulus and slowest when incorrectly predicting this stimulus.

Whitman and Geller (1971a, b, 1972) considered a continuous rather than a dichotomous expectancy model to account for sequential effects of prediction outcome on choice RT. Specifically, to interpret the result that choice reactions to correctly predicted events were faster when the preceding prediction outcome was correct rather than incorrect, Whitman and Geller proposed that correctly predicting a stimulus on a particular trial increased Ss' degree of expectancy for the next stimulus predicted. The demonstration that reactions following incorrect predictions in a choice RT task with compatible S-R mapping were slower when the preceding prediction outcome had been correct rather than incorrect (Whitman & Geller, 1971b) was consistent with such an expectancy model, since Ss' preparation for the occurrence of one stimulus alternative was presumed to be an inverse function of Ss' degree of expectancy for the other stimulus. To explain how the outcome of a prediction can influence stimulus expectancy, Whitman and Geller included confidence as an intervening variable and hypothesized that correctly predicting a stimulus presentation increases Ss' confidence in the next prediction, which in turn augments Ss' readiness for the predicted stimulus. The postulate that correctly predicting the less probable

*For their assistance in collecting data, the authors thank Ernie L. Viers, Mark Lewis, and Glenda J. McGuire. Portions of this research were presented at the meeting of the Psychonomic Society, St. Louis, November 1972.

†Requests for reprints should be sent to E. Scott Geller, Department of Psychology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061.

stimulus effects a greater increase in S's confidence for a subsequent prediction than does correctly predicting the more probable stimulus was adopted in order to interpret a reliable interaction between preceding stimulus probability and preceding prediction outcome (Whitman & Geller, 1972).

In order to study the validity of a confidence construct to explain changes in stimulus expectancy, the present experiment was designed to derive relationships between choice RT and the degree of stated confidence associated with a stimulus prediction. According to the theorizing of Whitman and Geller (1971a, b, 1972), choice RT should be an inverse function of the degree of confidence in a prediction when the predicted stimulus occurs but a direct function of confidence when the nonpredicted alternative is presented. In a previous two-alternative RT experiment, reaction speed was not significantly influenced by Ss' confidence judgments in their stimulus predictions (Geller & Pitz, 1970). However, it is possible that the complexity of the task in that study prevented the development of reliable confidence estimates sufficient to affect choice RT. In particular, each stimulus sequence consisted of only 10 presentations and the probability of each stimulus fluctuated intrasequentially as determined by E sampling one of two populations, the nature of the population being unknown to S. Also, the use of a scale and a movable pointer to indicate sequential confidence judgments might have added unnecessary complexity as well as contributing to commitment bias (cf. Geller & Pitz, 1968).

In the present RT experiments, a stimulus prediction and a confidence judgment in the prediction occurred prior to each presentation of one of two equiprobable stimulus alternatives. In the first experiment, the method of obtaining stimulus predictions and confidence judgments was varied among Ss. One group of Ss made a verbal prediction and a confidence estimate prior to each stimulus presentation. Another group of Ss gave verbal predictions but E verbalized the degree of confidence associated with each prediction. For a third condition, E verbalized both a stimulus prediction and a confidence estimate.

EXPERIMENT I

Method

Subjects

Sixty Ss (36 males, 24 females) were assigned randomly to one of three experimental conditions, with the restriction that each group had the same sex distribution (12 males, 8 females). The Ss were volunteers from the introductory psychology classes at Virginia Polytechnic Institute and State University, and they had never participated in a RT experiment.

Procedure

During the individual 2-h sessions, Ss received 600 presentations of a .5-sec warning buzzer, followed by a random

time interval of between .5 and .2 sec and the symbol $\square$ or $\sqcap$. The stimulus alternatives were equiprobable and were displayed by the illumination of a 2 x 5 cm digital readout (Allied No. MSM5A-6007711), which was positioned in the center of a panel separating S from E. The 50-50 distribution of events was determined by filtering a uniform random-number generator on an IBM 360 computer system. The Ss' reaction task was to press one telegraph key following one of the equiprobable stimuli and another key following the alternative. The two response keys were located on a desk before S and were spatially separated by 25 cm. Ss positioned their left forefinger on the left key and their right forefinger on the right key. Half of the Ss in each experimental group pressed the left key to $\square$ and the right key following $\sqcap$. For the remaining Ss, the reverse S-R mapping arrangement was in effect. A choice reaction turned off the stimulus and stopped a digital millisecond timer. Anticipatory responses and incorrect choices were indicated to E via lights and reported to S immediately after they occurred.

The task instructions emphasized the importance of both speed and accuracy when making choice reactions. Ss were told that the sequence of 600 events was preprogrammed on punched paper tape but were given no information regarding the relative frequency of the stimulus alternatives. The E presented 10 practice trials at the beginning of the session and during that time answered S's questions concerning the task. Halfway through each session (i.e., after 300 RT trials), Ss were given a 10-min break period.

Group S-Predict/S-Confidence (S-P/S-C)

Prior to each stimulus presentation, Ss verbally predicted which stimulus they expected to occur ($\square$ labeled *up* and $\sqcap$ labeled *down*) and then indicated their confidence in the stimulus prediction by verbalizing the number 50, 60, 70, 80, 90, or 100. These confidence values were printed on a card that hung above the stimulus readout. The Ss were instructed that the confidence number should represent the judged probability of their prediction being correct. The E referred to the example that a judgment of 50 would represent complete uncertainty in the prediction, whereas a judgment of 100 would indicate complete certainty that the predicted stimulus would occur. Ss were told that each prediction and confidence estimate should represent a subjective expectancy and thus were encouraged to contemplate each judgment.

Group E-Predict/E-Confidence (E-P/E-C)

Ss did not give predictions and confidence estimates; instead, E provided a stimulus prediction and a confidence value prior to each stimulus presentation. The Ss were instructed that the confidence judgment would range from 50 to 100 in steps of 10 and that the number would represent the actual probability that the predicted event would occur. As for Group S-P/S-C, the six possible confidence values were displayed above the stimulus readout. The task instructions emphasized that a judgment of 100 would indicate that the predicted event would definitely occur and that, for judgments of 70, the predicted event would actually occur on 70% of the trials, while, following 30% of such judgments, the predicted stimulus would not appear.

These Ss received the same sequence of events as Ss in Group S-P/S-C. Each of the six confidence values occurred an equal number of times in the sequence (i.e., 100), and the sequential arrangement of the predictions and judgments was derived by filtering a random-number generator on an IBM 360 computer. The predictions for the series of 600 events were selected so that the probability of the prediction being correct was exactly as indicated by the confidence estimate, and the percentage of correct and incorrect prediction outcomes was balanced between the two stimulus alternatives. For example, of the 100 confidence judgments of 80, 80% of the predictions were correct (40% for *up* and 40% for *down* predictions) and

20% were incorrect (10% *up* and 10% *down* predictions); for the estimates of 50, half of the predictions were incorrect, 25% for predictions of *up* and 25% for *down*.

Group S-Predict/E-Confidence (S-P/E-C)

The Ss received the same stimulus sequence as those in Groups S-P/S-C and E-P/E-C. Prior to the occurrence of each stimulus event, S verbalized which stimulus alternative he expected and then E indicated the degree of confidence S should have in the prediction. The confidence numbers that were displayed above the stimulus readout varied from 0 to 100 in steps of 10 and represented the actual probability that the predicted stimulus would occur. On a particular trial, when S predicted the same stimulus that E predicted in Condition E-P/E-C, the confidence value was identical to that given Ss of Group E-P/E-C and ranged from 50 to 100. However, when the stimulus predicted was the alternative to that predicted by E in Condition E-P/E-C, the confidence judgment was determined by subtracting the judgment of that trial from 100. For example, if the judgment for a stimulus on a given trial was 90 for Group E-P/E-C and S predicted the other stimulus, then E gave the judgment of 10.

Results

Inaccurate choice reactions (stimulus anticipations and inappropriate identifications) were eliminated from the data analysis and did not reach 5% (i.e., 30) for the Ss reported. One S in Group S-P/S-C made 63 errors (i.e., > 10%) and his data were disregarded. The correct latencies of each S were dichotomized into correct and incorrect prediction outcome categories and then classified according to the associated confidence (C) value. For Group E-P/E-C, the C-100 category consisted of only correct predictions. Since the C categories of Groups S-P/S-C and S-P/E-C were S controlled, categories with a small number of data entries sometimes occurred. Six Ss in Group S-P/S-C and three Ss in Group S-P/E-C gave less than four judgments of 50 or 100. Four of these nine Ss never indicated complete uncertainty (i.e., 50) in a stimulus prediction. Dichotomizing Ss' C values according to prediction outcome resulted in at least one category with less than two data entries for these nine Ss; therefore, these Ss were replaced. The consequence was that each S contributed at least five latencies for each prediction-C by prediction outcome category. The category means were calculated for each S and group averages were derived.

An overall analysis of variance (ANOVA) compared RT as a function of prediction outcome and the analogous C values of the three experimental conditions (i.e., Cs of 50, 60, 70, 80, and 90) and was, therefore, a factorial of 3 (groups) by 2 (prediction outcomes) by 5 (C values). Each main effect and interaction of this analysis was significant (all p s < .05). The subsequent analyses which resulted from each of the three possible two-group comparisons used the appropriate error terms from this overall ANOVA.

Fig. 1. Choice reaction time as a function of prediction outcome and confidence in the stimulus prediction when S made predictions and confidence judgments (Group S-P/S-C) and when E gave the predictions and the confidence estimates (Group E-P/E-C).

Group S-P/S-C vs Group E-P/E-C

Figure 1 depicts the mean latency for each of the 12 categories of Group S-P/S-C and 10 of the 11 possible categories of Group E-P/E-C. The mean latency for the C-100 category of Group E-P/E-C was 357 msec and is not depicted. For both experimental conditions, choice RTs were markedly shorter following correct than incorrect predictions; given a correct prediction, choice RT was an inverse function of C. When the predicted stimulus did not occur, the relationship between C and RT was a linear increasing function for Group E-P/E-C, but for Group S-P/S-C the function was not direct but slightly inverse.

Disregarding the C-100 categories, the ANOVA was a 2 (group) by 2 (prediction outcome) by 5 (C) factorial. Significant main effects of group [$F(1,57) = 6.17$, $p < .025$], prediction outcome [$F(1,57) = 95.78$, $p < .001$], and C [$F(4,228) = 7.50$, $p < .001$] were obtained. All interactions were significant, i.e., Group by Prediction Outcome [$F(1,57) = 4.89$, $p < .05$], Group by C [$F(4,228) = 6.37$, $p < .001$], Prediction Outcome by C [$F(4,228) = 10.45$, $p < .001$], and Group by Prediction Outcome by C [$F(4,228) = 2.41$, $p < .05$].

Fig. 2. Choice reaction time as a function of prediction outcome and confidence in the prediction when E gave the predictions and the confidence judgments (Group E-P/E-C) and when S made the predictions but E indicated the prediction confidence (Group S-P/E-C).

The separate ANOVA for Group S-P/S-C included the C-100 category and indicated main effects of prediction outcome and C ($ps < .001$), and a Prediction Outcome by C interaction ($p < .005$). For this group, the simple effects analyses of choice RT as a function of C for correct and for incorrect predictions indicated a significant effect of C when the prediction was correct ($p < .001$) but not when the prediction was incorrect ($.25 > p > .10$).

The 2 (prediction outcome) by 5 (C) ANOVA for Group E-P/E-C indicated a significant main effect of only prediction outcome and a Prediction Outcome by C interaction ($ps < .001$). The simple effects analyses indicated reliable main effects of C for both correct and incorrect prediction outcomes ($ps < .001$).

As shown in Fig. 1, the main effect of group and the Group by C interaction obtained with the overall ANOVA were largely due to differential effects of group and C on RT following incorrect predictions. In fact, the ANOVA for the incorrect prediction categories indicated reliable effects of only group [$F(1,57) = 4.94, p < .05$] and the Group by C interaction [$F(4,228) = 4.34, p < .005$]; whereas the ANOVA for correctly predicted events did not result in a significant group effect ($p > .10$) but did indicate a main effect of C [$F(4,228) = 14.57, p < .001$]. The Group by C interaction was also significant for this analysis

[$F(4,228) = 3.40, p < .01$] and was largely due to the reliable group difference for correct predictions in the C-50 or "complete uncertainty" category (Tukey's HSD = 56.8, $p < .01$).

Group E-P/E-C vs Group S-P/E-C

For both groups, Ss reacted faster to correctly predicted events than to incorrectly predicted events; however, comparing the analogous C categories between groups, the latency difference between correct and incorrect predictions was consistently larger when Ss gave the predictions (i.e., Group S-P/E-C). When the C value given by E for Group S-P/E-C was above rather than below 50 and implied that S should expect the predicted rather than the nonpredicted stimulus, the latency difference due to prediction outcomes was more prominent. The ANOVA for choice RT as a function of the analogous C categories of the two groups (i.e., 50 to 90) indicated a significant main effect of prediction outcome [$F(1,57) = 123.36, p < .001$] and interactions of Group by Prediction Outcome [$F(1,57) = 10.84, p < .005$] and C by Prediction Outcome [$F(4,228) = 27.01, p < .001$]. No other effects reached significance at a .05 level.

The ANOVA that compared the 10 latency categories for correct and incorrect predictions of Group S-P/E-C, i.e., a 2 (prediction outcome) by 10 (C) factorial, indicated significant main effects of prediction outcome [$F(1,19) = 26.01, p < .001$] and C [$F(9,171) = 12.68, p < .001$]. In addition, a significant Prediction Outcome by C interaction was obtained [$F(9,171) = 105.58, p < .001$]. Although Ss' latencies were shorter to incorrect than to correct stimulus predictions when E's C estimate for the predicted event was less than 50, the main effect of prediction outcome demonstrated that Ss responded reliably faster when the C value verbalized by E referred to the stimulus presented than when the stimulus was not correctly predicted by S and the corresponding expectancy value for the presented stimulus could be inferred as 1-C. The main effect of C was due to a monotonic decreasing relationship between RT and C when the prediction outcome effect was not considered.

Group S-P/S-C vs S-P/E-C

The ANOVA of RT as a function of prediction outcome and the five analogous C categories of these groups demonstrated significant main effects of group [$F(1,57) = 5.20, p < .05$], prediction outcome [$F(2,57) = 171.08, p < .001$], and C [$F(4,228) = 8.35, p < .001$]. Of the interactions, only the Group by Prediction Outcome was *not* significant ($p > .10$), i.e., Group by C [$F(4,228) = 5.64, p < .001$], Prediction Outcome by C [$F(4,228) = 21.33, p < .001$], and Group by C by Prediction Outcome [$F(4,228) = 3.94, p < .005$]. The 2 (groups) by 5 (C) analysis for correct

predictions demonstrated a reliable effect of only C [$F(4,228) = 20.97$, $p < .001$], while the ANOVA for incorrect predictions indicated significance for only the Group by C interaction [$F(4,228) = 7.02$, $p < .001$]. The Tukey HSD test indicated that the prominent group difference between latencies to correct predictions of the C-50 category was reliable ($p < .01$).

Discussion

Unlike the research of Geller and Pitz (1970), the present study found confidence in a stimulus prediction to be a significant determinant of choice RT. When E indicated the confidence judgment on each trial (i.e., Groups E-P/E-C and S-P/E-C), the relationship between RT and confidence was exactly as hypothesized by the continuous expectancy notions used previously to explain the independent influence of stimulus probability and prediction outcome on two-choice RT (Geller & Pitz, 1970; Geller et al, 1971, 1972) and to interpret the sequential effects of prediction outcome on choice RT (Whitman & Geller, 1971a, b, 1972). In contrast to RT models which presume that stimulus expectancy varies on each trial in an all or none fashion (e.g., Falmagne, 1965; Theios & Smith, 1972), the continuous expectancy model proposes that Ss' readiness to identify a stimulus may vary continuously and that a certain degree of preparedness for one stimulus results in the same amount of unpreparedness for the alternative stimulus. Although the present finding that choice RT varied significantly as a function of confidence within both correct and incorrect prediction outcome categories was clearly hypothesized by a continuous expectancy theory, it should be noted that such results do not refute a two-state RT model. That is, the present results may be explained by the notions that only a "set" or "unset" expectancy state exists on a given trial and that the probability of S being in the prepared state increases directly with the degree of S's confidence in the preparation (i.e., prediction).

When Ss verbalized both the prediction and the confidence estimate prior to each presentation, choice RT to incorrectly predicted stimuli was not significantly influenced by Ss' confidence judgment. This finding is not consistent with the premise that the degree of preparedness for one stimulus implies an analogous degree of nonpreparedness for the alternative. Correspondingly, the results of some other RT studies suggest that in some situations a continuous expectancy model may only apply to reactions following correctly predicted stimuli. For example, Whitman and Geller (1973) manipulated the probability of a correct prediction between Ss and found RT to be an inverse function of the prediction probabilities .10, .30, .50, .70, and .90 if S's stimulus prediction was correct, but RT was not reliably influenced by the probability of a correct prediction when the nonpredicted stimulus was

presented. Presuming that prediction probability and prediction confidence are directly related, these relationships observed by Whitman and Geller are consistent with those for Group S-P/S-C of the present study. In addition, a correct prediction on a given RT trial (assumed to effect an increase in subsequent prediction confidence) was always found to facilitate subsequent reactions to correctly predicted events in comparison with reactions to correctly predicted events preceded by incorrect prediction outcomes (Whitman & Geller, 1971a, b, 1972). On the other hand, the effects of preceding prediction outcome on reactions to incorrectly predicted stimuli were neither robust nor consistent. According to the continuous expectancy model, reactions to incorrect predictions should be inhibited to a greater extent when the previous outcomes are correct rather than incorrect. However, this hypothesized sequential influence of prediction outcome was observed in only one RT experiment and the effect was slight (cf. Whitman & Geller, 1971b).

The experimental condition in the present study not only affected the relationship between confidence and RT to nonpredicted events but also determined the extent of the prediction outcome effect. Specifically, regardless of the source of the confidence judgments (S or E), the latency difference between responses to correct and incorrect prediction outcomes was consistently greater when S rather than E indicated the stimulus predictions. For the comparison between Groups S-P/E-C and E-P/E-C, a greater expectancy due to personal predictions apparently augmented both response facilitation to identify correctly predicted events and response inhibition to identify incorrectly predicted events. Also, when both predictions and confidence judgments were given by S rather than E, reactions were prominently slower to incorrectly predicted events. However, following correctly predicted stimuli, reactions were *not* reliably faster when S rather than E made the predictions and confidence judgments. In fact, response latencies in the correct prediction, confidence-50 category were significantly slower for Group S-P/S-C than for Groups E-P/E-C and S-P/E-C; this finding implies that Ss were more uncertain in a stimulus prediction when they gave the confidence judgment of complete uncertainty.

Thus, the relationships between choice RT and prediction confidence were essentially the same for each experimental condition, except for the observation that RTs to incorrectly predicted stimuli were not significantly influenced by confidence when both predictions and confidence judgments were made by S. The extent to which this difference may be attributed to only the source of the confidence statements cannot be established in the present study, because the validity of the estimates also varied between Group S-P/S-C and the other conditions. Specifically, E's confidence ratings reflected the actual probability that the predicted stimulus would occur, but S's prediction confidence was

Fig. 3. Choice reaction time as a function of the outcome of S's predictions and E's random or true confidence judgments.

not related to the probability of being correct. Perhaps the differential effects of confidence on RT to incorrectly predicted events was due to Ss' placing more assurance in E's probability related confidence statements than in their own. Experiment II was designed to determine if the relationship between CRT and prediction confidence may be significantly affected by the degree to which a confidence statement is related to the probability of being correct. The E indicated a confidence estimate following each of Ss' predictions. The judgments were given randomly for one experimental condition; other Ss received estimates representing the actual probability that their predictions were correct.

EXPERIMENT II

Method

Subjects

Forty volunteers from introductory psychology classes at Virginia Polytechnic Institute and State University were assigned randomly to one of two experimental groups, with the restriction that the same proportion of males and females participated in each group (i.e., 12 males, 8 females).

Procedure

The apparatus was identical to that used in Experiment I. Each of the 10 practice and 300 experimental trials of the 1-h sessions began when S predicted which of the two stimuli would occur by verbalizing "up" to indicate $\square$ or "down" to predict $\square$.

Then E stated a confidence judgment for S's prediction. A .5-sec warning buzzer preceded the stimulus presentation by a variable time interval of between .5 and 2 sec. By pressing a left- or right-hand telegraph key, S made a choice response which then turned off the stimulus and stopped a digital millisecond timer. The S-R mapping relationship was balanced within the two experimental conditions.

Group Random

Following Ss' prediction, E verbalized one of three confidence estimates, 20, 50, or 80, which were printed on a card displayed above the stimulus readout. The three confidence levels were equiprobable and, therefore, each occurred on 50 trials with each stimulus. The sequence of confidence estimates was determined by a random-number generator on a computer, such that, regardless of the confidence estimate given, the probability of the prediction being correct was chance or .50.

Group True

As in Group Random, E verbalized one of the three confidence estimates following S's prediction, with the exception that the confidence indicated was the actual probability that S's prediction would be correct. That is, the confidence given by E on a particular trial was a function of S's prediction, the confidence sequence used for Group Random, and a "prediction" list which was constructed so that a true confidence estimate could be determined on each trial. Specifically, the "prediction" list was determined by the random confidence sequence and the stimulus sequence, such that the "prediction" list was identical to the stimulus sequence on 20% of the trials when the confidence was 20, on 50% of the trials when the confidence was 50, and on 80% of the trials when the confidence was 80. Thus, when Ss' prediction matched the "prediction" list, the confidence list estimate (C) was true and verbalized by E; when Ss' prediction was the alternative to the stimulus of the "prediction" list, the true confidence presented to S was $1 - C$. Consequently, the confidence estimate of 50 occurred for Ss of Group True on the same 100 trials as for Ss of Group Random. However, an estimate of 20 or 80 was dependent on Ss' prediction and, thus, the occurrence of these confidence estimates could not be controlled.

Results and Discussion

Trials on which anticipatory responses or incorrect choice reactions occurred were eliminated from the data analysis and did not exceed 5% (i.e., 15) for any S. The remaining choice RTs were categorized according to the prediction outcome (correct or incorrect) and the confidence level (20, 50, or 80). Category means were computed for each S and used to determine the group averages depicted in Fig. 3.

For both Group Random and Group True, choice RT to correctly predicted stimuli was a decreasing function of the confidence value. However, RT to incorrectly predicted stimuli was differentially influenced by the experimental conditions: Group True exhibited an increase in RT of 21 msec as a function of confidence, while Group Random showed a corresponding increase of only 6 msec. A 2 (group) by 2 (prediction outcome) by 3 (confidence) analysis of variance (ANOVA) indicated no main effect of group ($F < 1$), but reliable effects of prediction outcome [$F(1,38) = 34.91$, $p < .001$] and confidence [$F(2,76) = 4.12$, $p < .025$].

Of the interactions, two were *not* significant: Group by Prediction Outcome [$F(1,38) = 3.44$, $p > .05$] and Group by Confidence ($F < 1$), while two were reliable: Confidence by Prediction Outcome [$F(2,76) = 19.15$, $p < .001$] and Group by Confidence by Prediction Outcome [$F(2,76) = 3.66$, $p < .05$]. To investigate the possibility that the observed effects of confidence and prediction outcome were biased by differential changes in RT as a function of trials, an ANOVA included two trial blocks of 150 trials in addition to groups, prediction outcome, and confidence. A main effect of trials was reliable [$F(1,38) = 26.64$, $p < .001$], but none of the interactions involving trials were significant ($ps > .10$).

The simple effect of confidence was determined independently for RT to correctly and incorrectly predicted stimuli for Groups True and Random by computing F tests for each of the four curves depicted in Fig. 1. For Group True, the effect of confidence on RT was reliable, regardless of the prediction outcome: $F(2,76) = 17.60$, $p < .001$ and $F(2,76) = 4.39$, $p < .025$ for correct and incorrect prediction outcomes, respectively. In contrast, the simple effects tests for Group Random indicated a significant effect of confidence on RT to correctly predicted stimuli [$F(2,76) = 3.48$, $p < .05$] but not on RT to incorrectly predicted stimuli ($F < 1$).

Thus, a true confidence estimate reliably influenced RT to both correctly and incorrectly predicted stimuli, while a random confidence estimate significantly influenced RT to correctly predicted stimuli but not to incorrectly predicted stimuli. These observations suggest that the lack of an effect of prediction confidence on RT to incorrectly predicted stimuli for Group S-P/S-C of Experiment I was due to the between-group variability of the validity rather than to the source of the confidence judgment. The result that trial block was not involved in any significant interaction in Experiment II implied that Ss discriminated between true and random confidence statements relatively early in the event sequence. It is noteworthy that the results of both experiments indicated that differential assurance in confidence judgments between groups was only reflected by choice RT when Ss' stimulus predictions were incorrect.

In summary, the within-S effects of prediction outcome and confidence on choice RT observed in the present research suggested that S may not be only set or unset for a particular stimulus but that Ss' expectancy for a stimulus event may vary in degree within a set or an unset state. Not only did Ss react markedly faster to correctly predicted than to incorrectly predicted stimuli, but when S was prepared for the presented stimulus choice RT was an inverse function of the confidence estimate in the prediction. When choice RT to a nonpredicted stimulus was significantly influenced by the confidence that the predicted stimulus would occur,

the relationship was monotonically increasing. The between-group comparisons of Experiment I demonstrated that the source of the stimulus prediction influenced the extent to which choice RT was affected by the outcome of the stimulus prediction and that the source of the confidence judgment determined the relationship between confidence and RT to incorrectly predicted events and the degree to which reactions to correctly predicted stimuli were inhibited by a prediction confidence of complete uncertainty. The between-group comparisons in Experiment II showed that the validity of E's confidence estimate for S's prediction was sufficient to alter the effect of confidence value on choice RT to incorrectly predicted stimuli.

REFERENCES

- Bernstein, I. H., & Reese, C. Behavioral hypotheses and choice reaction time. *Psychonomic Science*, 1965, 3, 259-260.
- DeKlerk, L. F. W., & Oppe, S. Subjective probability and choice reaction time. In A. F. Sanders (Ed.), *Attention and performance III*. Acta Psychologica, 1970, 33, 243-251.
- Falmagne, J. C. Stochastic models for choice reaction time with applications to experimental results. *Journal of Mathematical Psychology*, 1965, 2, 77-124.
- Geller, E. S., & Pitz, G. F. Confidence and decision speed in the revision of opinion. *Organizational Behavior & Human Performance*, 1968, 3, 190-201.
- Geller, E. S., & Pitz, G. F. Effects of prediction, probability, and run length on choice reaction speed. *Journal of Experimental Psychology*, 1970, 84, 361-367.
- Geller, E. S., Whitman, C. P., & Farris, J. C. Probability discrimination indicated by stimulus predictions and reaction speed: Effects of S-R compatibility. *Journal of Experimental Psychology*, 1972, 93, 404-409.
- Geller, E. S., Whitman, C. P., Wrenn, R. F., & Shipley, W. G. Expectancy and discrete reaction time in a probability reversal design. *Journal of Experimental Psychology*, 1971, 90, 113-119.
- Hale, D. J. Sequential effects in a two-choice serial reaction task. *Quarterly Journal of Experimental Psychology*, 1967, 19, 133-141.
- Hinrichs, J. V., & Craft, J. L. Verbal expectancy and probability in two-choice reaction time. *Journal of Experimental Psychology*, 1971, 88, 367-371.
- Kanarick, A. F. Probability and payoff as factors influencing two choice reaction-time. *Journal of Engineering Psychology*, 1966, 5, 37-46.
- Keele, S. W. Repetition effect: A memory-dependent process. *Journal of Experimental Psychology*, 1969, 80, 243-248.
- Laming, D. R. J. Subjective probability in choice-reaction experiments. *Journal of Mathematical Psychology*, 1969, 6, 81-120.
- Remington, R. J. Analysis of sequential effects in choice reaction times. *Journal of Experimental Psychology*, 1969, 82, 250-257.
- Theios, J., & Smith, P. G. Can a two-state model account for two-choice reaction-time data? *Psychological Review*, 1972, 79, 172-177.
- Whitman, C. P., & Geller, E. S. Runs of correct and incorrect predictions as determinants of choice reaction time. *Psychonomic Science*, 1971a, 23, 421-423.
- Whitman, C. P., & Geller, E. S. Prediction outcome, S-R

- compatibility, and choice reaction time. *Journal of Experimental Psychology*, 1971b, 91, 299-304.
- Whitman, C. P., & Geller, E. S. Sequential effects of stimulus probability and prediction outcome on choice reaction time. *Journal of Experimental Psychology*, 1972, 93, 373-378.
- Whitman, C. P., & Geller, E. S. Prediction outcome distributions as determinants of choice reaction time. *Perception & Psychophysics*, 1973, 13, 105-107.
- Williams, J. A. Sequential effects in disjunctive reaction time: Implications for decision models. *Journal of Experimental Psychology*, 1966, 71, 665-672.

(Received for publication March 5, 1973;
accepted March 13, 1973.)