

CONTEXTUAL CONTROL OF ACCRUAL THRESHOLDS

William M. Petrusic and Joseph V. Baranski

Department of Psychology, Carleton University, Ottawa, Ontario, Canada,
bill_petrusic@carleton.ca;

Judgment and Decision Making Group, Defence Research and Development Canada,
Toronto, Ontario, Canada, Joe.Baranski@drdc-rddc.gc.ca

Current theories of decision processing in a wide variety of binary choice tasks posit some form of evidence accumulation over time until a threshold or criterion is reached. In the context of discrimination of visual extent, we show that evidence accrual criteria for any particular stimulus pair are dependent upon the overall, global, difficulty context. In particular, response times (RTs) on a target set of stimulus pairs of moderate difficulty were increased when embedded in a difficult context. Moreover, when easy to discriminate pairs were included along with difficult pairs, RTs were the same as when in a difficult context alone, consistent with the additional finding that target pair RTs were uninfluenced by the inclusion of easy pairs. Thus, the most difficult stimulus pair encountered over the course of the experiment controls evidence accrual criteria. The absence of contextual effects on either discriminative accuracy or confidence ratings is also entirely consistent with the Slow and Fast Guessing theory (Petrusic, 1992) based views of contextual difficulty effects.

In all classes of evidence accrual models of comparative judgement, decisional criteria are broadly set by the particular demands for speed versus accuracy and by various strategic biasing manipulations leading to a preference for one response alternative over the other. In addition, Vickers (1979) surmised that the global difficulty context might well determine how the accrual criteria might be set for any particular pair and he developed his adaptive accumulator module to account for how the presence of other alternatives might determine the setting of accrual criteria.

As well, Vickers provided a detailed theoretical analysis of global difficulty contextual difficulty effects in terms of his adaptive module. On his view, the primary decisional criterion, k , was regulated by a secondary adaptive process. On each pass through the accrual process the current level of confidence is compared to a preset target level of confidence and the discrepancies are accumulated in competing “under confidence” and “overconfidence” accumulators. When, for example, the difficulty context is increased, and the consequent amount of accumulated under confidence reaches its preset criterion, this triggers an upward adjustment in the primary decisional criterion, k . Vickers (1979, p. 203) provides the following example of such a dynamic adjustment to the global difficulty context.

...if a stimulus difference of size m were embedded within a series of very small differences, we should normally expect an observer to make a slower, more careful judgment of m than if it were embedded in a series of very large differences, and we would “explain” this by arguing that the observer would be inclined to adopt higher values of k when the task was generally more difficult.

On Vickers' view, and indeed, for all evidence accrual models, an increase in the primary decisional criterion will result in an increase in RTs and discriminative accuracy; i.e., a *speed-accuracy tradeoff*.

Petrušić and Baranski (1989) initiated the experimental study of the effects of variations in the overall difficulty context. Precisely as suggested by Vickers, in our first experiment, in one condition a target set of pairs (line lengths) was embedded in a difficult context and in an easy context in another condition. RTs to the target pairs were substantially and reliably larger when they were embedded in the difficult context than in the easy context. As well, Petrušić and Baranski (1997) provide a replication of these immediate global difficulty context effects.

With a view toward adding a more general perspective to these global contextual difficulty effects, the effect of context shift was examined under both accuracy and speed stress conditions. In the accuracy condition, the shift from the easy to the difficult context resulted in an increase in RTs. Moreover, as in the first experiment neither the accuracy of discrimination nor overall confidence were influenced by the effects global difficulty context. Interestingly, in the speed stress condition, RTs were unaffected by the difficulty context but discriminative accuracy decreased as did confidence with the shift to the more difficult context.

More recently Baranski and Petrušić (2003) replicated and extended our earlier results, primarily by making the shifts in context within sessions. RTs to target pairs increased with the shift to a more difficult context but neither discriminative accuracy nor confidence were affected. In a second experiment a shift from a hard to an easy context had no effect on any dependent measure. In a final experiment, with a shift from an easy to a difficult context, under speed stress, as in Petrušić and Baranski (1998), discriminative accuracy and confidence decreased.

Taken together, these findings provide limited support for the operation of Vickers' adaptive module. Notably, the increases in target RTs in the shift to the more difficult context under accuracy stress was not accompanied by an increase in discriminative accuracy. Second, the shift to the easy context should have resulted in faster target RTs and reduced accuracy on the target pairs. Finally, the decreased accuracy on target pairs under speed stress is not easily, if at all, permitted.

On the other hand, each of our findings is predicted in the context of Petrušić's (1992) Slow and Fast Guessing Theory (SFGT). SFGT, a variant on LaBerge's (1962) discrete accumulator, posits cut points, C_1 and C_2 on a decision axis, $d=X-Y$, such that evidence favouring the response R_1 occurs if $d < C_1$ with probability p_1 and evidence favouring the response R_2 occurs if $d > C_2$ with probability p_2 and evidence favouring a state of doubt and occurs if $C_1 < d < C_2$ with probability $p_3 = 1 - (p_1 + p_2)$. We refer to the interval $C_2 - C_1$ as the interval of uncertainty (IOU) and accruing a criterion amount of information falling in the IOU (or inconclusive information) triggers a state of doubt, or indifference, and guessing is the basis for responding. Indeed, letting \hat{a}_1 , \hat{a}_2 , and \hat{a}_3 denote evidence accrual thresholds for responses R_1 , R_2 , and the doubt state, respectively. Slow guessing arises when \hat{a}_3 is set high and fast guessing arises when \hat{a}_3 is set very low (e.g., 2).

In contrast to Vickers' account of adaptation to difficulty context, we assumed that an increase in the global judgment difficulty context leads participants to secure more diagnostic information. According to SFGT, this notion is conceptualized in terms of expanding the interval of uncertainty and implies that as judgments become more difficult we tend to base our decisions on more 'certain' information. SFGT-IOU view implements this idea through an increase in p_3 .

The present experiment was designed to provide further empirical scrutiny of these two accounts of how the global difficulty context controls the setting of decisional criteria.

The contextual manipulation employed a target set of moderately difficult to discriminate pairs of lines and two induction conditions, one hard and the other easy. A Baseline condition was defined by the target set of pairs, a Hard condition arose from adding the hard stimulus pair to the target set, an Easy condition arose from adding the easy pair to the target set, and a Mixed condition was defined by adding both the hard and the easy pair to the target set.

Vickers' adaptive module predicts increases in RTs and discriminative accuracy with target pairs in the Hard condition (relative to Baseline), and decreases in RTs and accuracy in the Easy condition, and no effects in the Mixed condition, on the assumption that criterion adjustments are symmetric with the hard and the easy context pairs. On the other hand, SFGT-IOU view is clear in predicting that RTs for target pairs in the Hard and Mixed condition will be increased with no changes in accuracy and the inclusion of the easy pair in the Easy condition will have no effect on target pair RTs or accuracy.

Method

Participants. Forty-two Carleton University undergraduate students participated for one experimental session of approximately 1.5 hours duration in return for course credit.

Apparatus. Stimuli were presented on a Zenith ZCM-1492 flat screen video monitor, 245 mm (640 pixels) horizontally and 185 mm (480 pixels) vertically. Timing, accurate to within ± 1 ms, was possible with a Data Translation clock board and extensive software development. Graphics production, stimulus presentation, event sequencing, and the recording of responses and response times were controlled by an IBM-PC 486DX, 33 MHz clone computer.

Participants responded on a movable 13 by 17 cm panel containing three banks of response buttons. The first at the bottom and center of the panel contained a single, home key, which was used to initiate a trial. The second directly above contained two primary response buttons that were used to indicate a left or a right decisional response. The third bank, located directly above and symmetrically around the second bank, contained six confidence rating buttons arranged in a semi-circle with labels "50", "60", "70", "80", "90", and "100" placed directly above the respective buttons from left to right.

Stimuli and design. The stimuli consisted of five pairs of horizontal visual extents presented side by side in the middle of the video monitor. Each stimulus pair is defined by the notation (x, y), where x denotes the length in pixels (2.61 pixels = 1 mm) of the line appearing on the left of the video monitor and y the length of the line on the right. The five stimulus pairs in their two presentation orders were (200, 200; 200, 200), (200, 202; 202, 200), (200, 206; 206, 200), (200, 210; 210, 200), and (200, 220; 220, 200) and these pairs are labelled Pair 1, Pair 2, Pair 3, Pair 4, and Pair 5, respectively. The five stimulus pairs are expressed, in terms of difficulty, by the ratios 1.00, 1.01, 1.03, 1.05, and 1.10, for Pairs 1 through 5, respectively. Pairs 1 and 5 serve as context induction pairs, defining extremely difficult and very easy contexts, respectively. Pairs, 2, 3, and 4 are target pairs used to gauge the possible contextual effects of the induction pairs.

Ten participants were randomly assigned to each of four groups. Participants in the Baseline condition made comparisons with just the target pairs. The Difficult context condition was defined by adding pair 1 to the set of target pairs and the Easy context condition arose upon adding pair 5 to the set of target pairs. Finally, the Mixed context condition was defined by adding both pair 1, the difficult pair, and pair 5, the easy pair, to the set of target pairs.

All participants received five blocks of 48 trials per block with target pairs. On half of the trials, participants selected the shorter line length in the pair and the longer on the other half. As well, each pair was presented in the two presentation orders. The 48 trials in each block with each pair arose from 4 replications of the factorial combination of the two

instructions, the two presentation orders, and the three pairs defining the target set. Participants in the Difficult context received an additional 16 trials on each block with pair 1 (4 replications of the factorial combination of the two instructions and the two presentation orders), and participants in the Easy context condition received 16 additional trials with Pair 5. Finally participants in the Mixed condition received a total of 80 trials in each block, the 48 with the three target pairs, 16 with pair 1 and 16 with pair 5. The first block of trials was preceded by 8 practice trials, randomly selected from the set defining a block for each group.

Procedure. Participants were instructed that following their primary decision they would be required to indicate how confident they were that the decision they had just rendered was in fact correct. They were told that the button labelled “50” represented a guess, “100” complete certainty, and that the remaining confidence categories denoted accordingly varying degrees of confidence.

Each trial was initiated with the participant depressing the home key, and 500 ms later one of the two instructions appeared and remained on the screen until the participant indicated his or her response. One s later the stimulus pair was presented. After the participant had made the primary decision, one of the six confidence keys was selected. The next trial began 2 s later. All participants were instructed that both speed and accuracy were important.

Results

For each participant, in all analyses, the dependent variables in each cell of the design are the mean RT for all responses and the mean percentage of errors. In each analysis of variance (ANOVA), Huynh-Feldt epsilon adjustment of degrees of freedom was used. However, the degrees of freedom associated with each value of F are defined by the design and the Mean Square Errors provided in the text are those given by the conventional degrees of freedom. Level of significance was set at 0.05 throughout. Two participants' overall performance was at or near 50% correct with each stimulus pair and they were replaced. In addition, RTs less than 200 ms and longer than 10 s were removed. This accounted for 240 of the overall 12,800 trials (1.87 %).

Response time analyses

As the plots in Figure 1 and Panel A of Figure 2 show, RTs (with the target pairs) are controlled by the most difficult stimulus pair in the stimulus set. RTs in the Mixed and the Hard conditions are nearly identical as are the RTs in the Baseline and Easy conditions. However, the mean RT of the Mixed and Hard conditions (2513.05 ms) is substantially longer than the mean RT of the Baseline and Easy conditions (2060.58 ms). An ANOVA with the three target stimulus pairs, five blocks, and two presentation orders as within participant factors, and the four context conditions as the between participant factor was conducted. An a priori comparison revealed the difference between the mean of the Hard and Mixed conditions differed reliably from the mean of the Baseline and Easy conditions ($F(1, 36)=7.65$, $MS(Error)=32119034.4$). The three stimulus pairs also differed, as expected, in terms of a priori difficulty of discrimination ($F(2, 72)=21.37$, $MS(Error)=503326.1$). The interaction between stimulus pair and presentation order was reliable ($F(2, 72)=3.41$, $MS(Error)=2957.4$), reflecting the fact that generally RTs were shorter in the short-long (shorter line on the left) presentation order than in the long-short presentation order but this form of space error was especially accentuated with the easiest of the target stimulus pairs. The main effect of block was also statistically reliable ($F(4, 144)=18.39$, $MS(Error)=931992.5$). As is evident in the plots in Figure 1, RTs decrease with increasing practice. Furthermore, as is also evident from the plots in Figure 1, the effects of practice are most evident with the Mixed and Hard

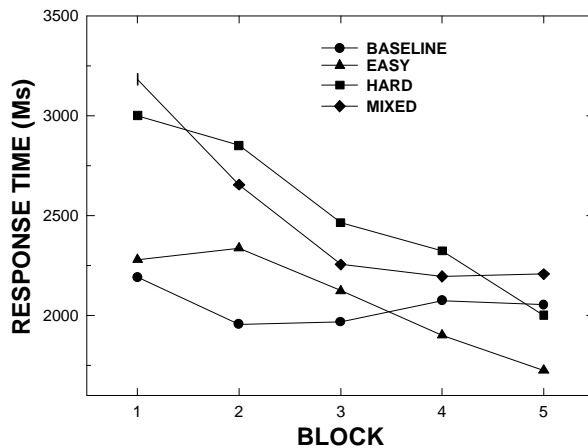


Figure 1. Mean RTs in each condition for each block.

conditions. Indeed by the fifth block the four context conditions no longer differ. This dependence of context condition on the amount of practice is reflected in a reliable interaction between the context condition and the block factor ($F(12, 144)=2.87$, $MS(Error)=931992.5$).

Discriminative accuracy

As is evident from the plots in Panel B of Figure 2, discriminative accuracy is uninfluenced by difficulty context ($F(3, 36)=1.18$, $MS(Error)=0.138$). Of course, as expected, the main effect of stimulus pairs in the target set was highly reliable ($F(2, 72)=254.32$, $MS(Error)=0.026$; $F(2, 72)=267.94$, $MS(Error)=0.141$ with arcsine transformed proportion correct); .579, .730, and .837 of the responses were correct for stimulus pairs 2, 3 and 4, respectively. The main effect of presentation order was also reliable ($F(1, 36)=4.94$, $MS(Error)=0.759$; $F(1, 36)=5.07$, $MS(Error)=4.20$ with arcsine transformed dependent measure). Discriminative accuracy was better for each of the three target pairs when the shorter line appeared on the left; i.e. the classic space error.

Confidence

As with discriminative accuracy, confidence was unaffected by context ($F < 1$). Also paralleling the discriminative accuracy measure, confidence reliably increased as the stimulus pair ratio increased ($F(2, 72)=56.91$, $MS(Error)=80.89$; i.e. as the a priori ease of

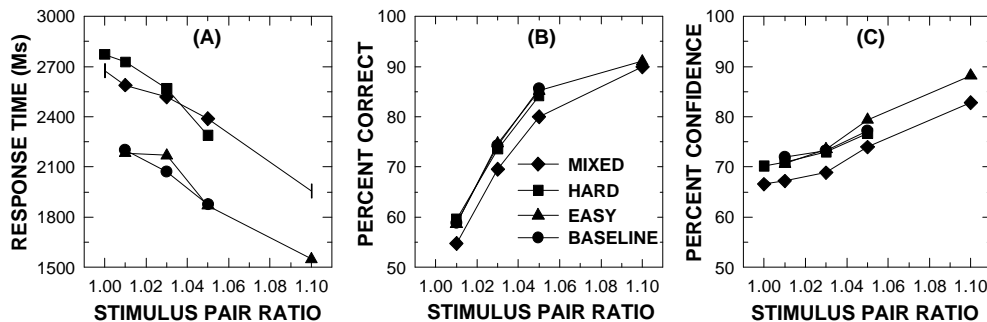


Figure 2. Mean RTs, Percent Correct, and Mean Percent Confidence with each stimulus pair in each condition.

discrimination increased. Confidence ratings also reflected the ubiquitous space error ($F(1, 36)=4.24$, $MS(\text{Error})=449.97$) with higher mean confidence ratings when the shorter line length appeared on the left. The main effect of presentation order was qualified by an interaction involving stimulus pair ($F(2, 72)=6.50$, $MS(\text{Error})=50.03$). Paralleling the effect of stimulus pair on presentation order with RTs, the space error was minimal with the most difficult pair (Pair 2) and was clearly evident with the remaining two pairs.

Discussion and Conclusions

The findings are clear, replicating and extending the Petrusic and Baranski (1989, 1997), global contextual difficulty effects and the Baranski and Petrusic (2003) within session difficulty shifts. The data are decisive in showing that the parameters of the decision process, for any particular stimulus pair, are controlled by the most difficult pair in the choice set. RTs to target pairs in the Mixed condition are the same as RTs in the Difficult condition. Moreover, the inclusion of easy pairs is of no consequence; RTs in the Easy context condition are the same as those in the Baseline condition. Baranski and Petrusic (2003) also showed that a shift from a difficult to an easy difficulty context had no effect on target pair RTs; only shifts to a more difficult context resulted in increases in target pair RTs. Together with the fact that neither discriminative accuracy nor confidence are influenced by global difficulty context is entirely in accord with the SFGT-IOU theoretical view. Namely, embedding target pairs within a set of more difficult comparisons results in an increase in the size of the IOU (i.e., an increase in the p_3 parameter) along with possible increases in the evidence accrual criteria of the discriminative states. As is also evident, the present findings provide little, if any, support, for Vickers' adaptive module.

Baranski and Petrusic (2003) noted that the increase in RTs with no change in accuracy leads one to question the rationality of the adaptive regulatory process. There is some evidence in the present experiment of movement toward rationality with increasing practice; the contextual difficulty induced increases in RTs were absent after extended practice.

References

Baranski, J.V. & Petrusic, W.M. (2003). Adaptive processes in perceptual comparisons: Effects of change in the global difficulty context. *Journal of Experimental Psychology: Human Perception and Performance*, **29**, 658-674.

Petrusic, W.M. & Baranski, J.V. (1989). Context, context shift, and semantic congruity effects in comparative judgments. In D. Vickers and P. L. Smith (Eds). *Human information processing: Measures, mechanisms and models*. Amsterdam: North-Holland.

Petrusic, W. M., & Baranski, J. V. (1997). Context, feedback, and the calibration and resolution of confidence in perceptual judgments. *American Journal of Psychology*, **110**, 543-572.

Vickers, D. (1979). *Decision processes in visual perception*. New York: Academic Press.

Acknowledgement

This work was supported by a Natural Sciences and Engineering Research Council of Canada Discovery grant to Petrusic.