The Impact of Complete and Selective Feedback in Static and Dynamic Multiple-Cue Judgment Tasks

Oren Griffiths (ogriffiths@psy.unsw.edu.au)
Ben R. Newell (ben.newell@unsw.edu.au)
School of Psychology, University of New South Wales, Sydney, Australia

Abstract

It is widely accepted that feedback is critical to guide the learning process and to make effective decisions. However, ideal feedback is relatively rare in our daily environments. Two multiple-cue judgment experiments examined whether biased and incomplete feedback leads to less accurate learning than comprehensive feedback. Experiment 1 found that participants given selective feedback produce equivalently accurate predictions, learned equally rapidly, and exhibited equivalent task structure knowledge to those given full feedback. Experiment 2 showed that selective feedback led to more accurate outcome predictions and task structure knowledge in a dynamic task environment. These results are problematic for error-based models of learning.

Keywords: Learning, decision making, dynamic environment, prediction error

Introduction

The importance of feedback in learning can scarcely be overstated. Feedback can be seen in many aspects of our lives: at work (audits, trial and error approach, quality control), while shopping (customer service evaluations), at home (reward/punishment of children) and, of course, in educational settings (exam marks, final grades).

Consequently, most formal learning theories suggest that appropriate feedback is necessary for learning (excluding models of unsupervised learning). For example, error-based models of associative learning (e.g. Rescorla & Wagner, 1972; Mackintosh, 1975), error-based models of categorization (e.g. Nosofsky, 1986; Kruschke, 1992), connectionist learning models based upon the delta-rule (Rumelhart & McClelland, 1986) and Bayesian models of learning/reasoning (for a review, see Oaksford & Chater, 2008) all suggest that learning only occurs when feedback is received. (We will hereafter cumulatively refer to these models as *error-based* models). By extension, these error-based models suggest that optimal learning is achieved when accurate, unbiased feedback is delivered as often as possible.

Selective Feedback

Unfortunately, optimal feedback conditions like these are rare outside of the laboratory (e.g. Hogarth, 2005). To illustrate, suppose your friend Susan wants to learn how to make good investments in the stock market. To do this, she picks 10 companies that she thinks will be profitable, invests in them and monitors the share prices of these companies over the following 6 months. However, she probably will not monitor the share prices of 10 companies that she did *not* invest in. Thus, the feedback that Susan receives will be both incomplete and biased. It is incomplete because she received feedback on only a portion of the relevant learning opportunities, and it is biased because the companies about which she received feedback were systematically related; they were the companies that she predicted to be profitable. What impact would this biased, selective feedback have upon Susan's learning? According to the aforementioned learning models it ought to be detrimental to her learning, relative to a scenario in which she was provided with balanced, comprehensive feedback.

Elwin, Juslin, Olsson & Enkvist (2007; see also Henriksson, Elwin & Juslin, 2008) examined this question using an experiment based on the prior example. On each trial participants were shown how well a particular company scored on four performance indices (e.g. market share, turnover; hereafter referred to as cues), which were each scored between 0 and 10. Participants were asked to predict how many percentage points profit (or loss) that company would make in the following year. Importantly the feedback provided differed between participants. After generating their predictions, the full feedback group was always told how much profit each company made. In contrast, the selective feedback group was only told about the profitability of companies that they predicted would be profitable. At test, the two groups did not differ in their abilities to predict the profit made by each company (if anything, a non-significant advantage was seen for those given selective feedback). That is, the extra feedback given to the full feedback group did not improve performance over those given selective feedback. This lack of a difference does not appear to be due to a ceiling or floor effect, as the mean prediction accuracy at test ranged from 0.75 – 0.9. At first glance, this result does not seem to be supportive of the standard, error-based view of learning.

However, there are alternative explanations of this result which are consistent with error-based learning. The selective group in the Elwin et al. (2007) study were given twice as many training trials as those in the full feedback group. This was done in order to equate the number of trials in which feedback was provided between the groups. However, it meant that from an error-based learning view, both groups experienced an equivalent number of learning opportunities (trials in which feedback was provided) making it less surprising that the groups performed equivalently at test.

Further, it is not clear to what extent participants in either group understood the task structure, such as the cue-weightings and the base-rate of profitability. It is possible that participants in the full feedback group had a richer understanding of the task structure but, for some reason, this understanding was not translated into more accurate prediction performance. In support of this suggestion, the full feedback group predicted profit at a rate that was closer to the objective base-rate of profitability (0.5) than the selective feedback group. Finally, Elwin et al. only reported prediction accuracy for the test trials. The benefit of full feedback may best be seen in the training phase (for example, the full feedback group may learn more rapidly).

Experiment 1

Experiment 1 was conducted to examine these alternative explanations of Elwin et al.'s data, and to seek a replication of their main findings. Four changes were made to Elwin et al.'s experiment in order to address these alternative explanations. First, the number of trials was equated for the full and selective feedback groups. This meant that the full feedback groups received up to twice as much feedback as the selective groups. Secondly, trial-by-trial training performance was analyzed in order to examine any differences in the rate at which learning occurred. Thirdly, at test, participants were asked to rate how important each cue was when generating their predictions (a measure of their knowledge of the cue-weightings). They were also asked to estimate the percentage of companies that were profitable in training (a measure of their knowledge of the base-rate of profitability).

Finally, two different base-rates were used. For half of the participants (the 50% groups), 50% of the companies shown in training were profitable, while for the remainder (the 80% groups) 80% were profitable. If the extra feedback allowed participants in the full feedback condition to learn more accurately about the task structure, including the base-rate, then they should be more sensitive to these different base-rates. In all other respects Experiment 1 was identical to that conducted by Elwin et al.

Method

Participants. Forty-eight introductory Psychology students from UNSW participated for course credit.

Design. Two between-subject manipulations were conducted. Half of the participants were given feedback on every trial (full feedback groups) and half were only given feedback on trials in which they predicted profit (selective feedback groups). Orthogonal to this manipulation, for half of the participants 50% of the companies shown in training were profitable, and for the remaining half, 80% were profitable. Participants were evenly and randomly allocated into these four groups.

Materials & Procedure. The materials and procedure were based upon those of Elwin et al. Participants were given a computerized profit prediction task. On each trial, participants were given four cues, which were the company's scores (0 – 10; higher scores indicate better performance) on each of four economic indices (market share, turnover, staff experience, expenses). Once shown the cues, participants made a prediction as to how many percentage points of profit (or loss) that company would make in the following year, on a scale ranging from 50% loss to 50% profit.

The profit made by each company was calculated by taking each cue value and multiplying it by the weight of that cue (1, 2, 3 or 4) and then summing these values. This yields a value between 0 and 100. For the 50% base-rate groups, 50 was then subtracted from this sum so as to make the company profits range from -50 (loss) to +50 (profit). For the 80% group, 35 was subtracted from the weighted sum of the cue-values. This yielded values between -35 and 65, in which approximately 80% of the values were positive. These values were then adjusted to fit the -50 to +50 scale (scores greater than 0 were multiplied by 50/65 and scores below 0 were multiplied by 50/35). The profit scores were calculated in this manner in order to (i) keep the profit range in the 80% groups equal to that for the 50% groups (-50 to +50), and (ii) to keep the distribution of the profit scores similar between the 80% and 50% groups.

After making each profit prediction, the full feedback groups were shown the correct answer. The selective feedback group, in contrast, were only shown the actual profit of that company if they had predicted the company to be profitable (i.e. had predicted a profit greater than 0). Otherwise these participants were simply moved to the next trial. All participants were given 4 blocks of 60 training trials (240 training trials in total). The transition between blocks was not signaled to participants, and participants were not provided with any breaks in training.

Upon completion of training, all participants were told that their ability to predict profitability was to be tested, and that they were to re-rate a small selection of the companies that they had previously seen. The test procedure was very similar to the training procedure, except that only 60 trials were given and no feedback was provided.

Immediately before the first test-trial, participants were asked to estimate the percentage of companies that were profitable in training (a measure of the base-rate) on a scale ranging from 0 – 100%. Then, immediately after completing the last test-trial, participants were asked to rate how important each cue had been when making their predictions, on a scale of 0 to 100 (a measure of cue-weights).

Results & Discussion

Prediction Data. The absolute value of the difference between the predicted profit and actual profit on each trial was calculated, and averaged across all of the trials in each block to yield participants' mean prediction error per trial. Low error scores indicate accurate predictions. Participants' prediction accuracies across training and test are shown in Figure 1.

Participants' prediction accuracy scores in training were analyzed using planned, orthogonal contrasts in a multivariate ANOVA. Type I error was controlled at .05. Averaged across the four training blocks, no significant differences in accuracy were observed between the 50% groups, and the 80% groups, $F(1,44) = 1.92$. Within each of these base-rate groups, no significant differences in prediction accuracy were observed between participants given full feedback and those given selective feedback; for the 50% groups, $F(1,44) = 2.45$; for the 80% groups, $F < 1$. Thus, selective feedback did not impair overall prediction accuracy during training.

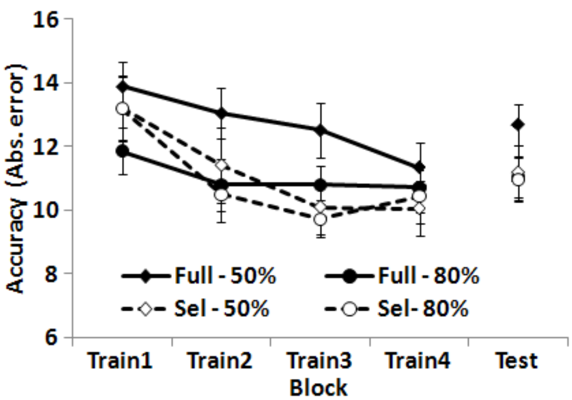

Figure 1: The mean prediction errors in Experiment 1. Error bars indicate the SEM for all figures.

To examine whether selective feedback impaired the learning rate, a linear trend analysis of participants' prediction error was conducted. Averaged across the four groups, a significant negative linear trend in prediction error was observed, $F(1,44) = 39.21$. This indicates that overall prediction accuracy increased across training. Amongst the 50% groups, participants given selective feedback showed a numerically larger decline in error across training than those given full-feedback, but this difference was not significant, $F < 1$. Similarly, amongst the 80% groups, the decline in error was numerically larger for the selective group, but again this difference was not significant, $F(1,44) = 2.47$. In summary, full feedback did not appear to benefit overall training accuracy or learning rate. If anything, the decline in error (a measure of learning-rate) was greater for the selective feedback groups, but these differences were not reliable.

The same between-subjects contrasts used to analyze participants' prediction accuracy in training were used to analyze their accuracy at test. No significant differences were observed, maximum $F(1,44) = 2.54$. Thus any differences in training performance were not transferred to the test phase.

Knowledge of task structure. Although no benefit for the full feedback groups was observed in participants' prediction accuracy (in training or test), these groups may have learned more about the task structure than the selective groups. To examine this prediction, participants' ratings of cue-importance and their estimates of the proportion of profitable companies were analyzed.

Participants' mean importance ratings for each cue are shown in Figure 2. A linear trend analysis shows that, averaged across the four groups, participants correctly rated the most heavily weighted cues as more important than the least weighted cues, $F(1,44) = 31.50$. This linear trend did not interact with the any of the between-subject, main effect contrasts, all Fs < 1. This suggests that neither the base-rate nor the feedback manipulation affected participants' knowledge of the cue-weightings.

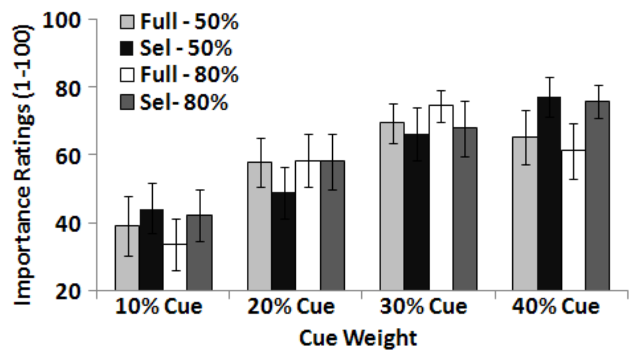

Figure 2: Mean ratings of cue importance in Experiment 1.

Participants mean estimates of the proportion of profitable companies were compared against the objective base-rate (either 50 or 80) with one-sample t-tests. Of the 50% base-rate groups, the full-feedback group significantly overestimated the base-rate, M = 58.42, $t(11) = 4.64$, but the selective feedback group's estimate did not significantly differ from the correct value, M = 51.17, $t(11) < 1$. Of the 80% groups, the mean estimate of the full feedback group (M = 74.25) was closer to the correct value than the selective feedback group (M = 71.75), but both groups significantly underestimated the base-rate, $t(11) = 2.21$and 3.68, respectively. Thus, there is mixed evidence regarding whether selective feedback impaired (or improved) participants' ability to monitor the base-rate of profitable companies.

In summary, the present data support and extend Elwin et al.'s findings. No evidence was found to suggest that selective feedback impaired learning or reduced participants' knowledge of the task structure, relative to participants given full feedback. Unlike Elwin et al.'s study, this cannot be due to uneven numbers of trials. The number of trials was equated between groups and thus, in the present study, the full feedback groups received more feedback than the selective groups. On average, the 50% and 80% base rate selective groups received feedback on only 61% and 80% of the trials, respectively. Nevertheless, no corresponding decrease in performance was observed.

One may wonder why no decrease was seen. Perhaps the present scenario was too simple to demonstrate the benefits of full feedback. For example, participants in the present task may have initially learned quite rapidly until they reached a performance level with which they were satisfied, and then ceased learning. On this account, the present study may not be sensitive to the influence of feedback because only a small amount of feedback may be required: enough to drive the initial learning process and then all further feedback could be ignored.

Experiment 2

To examine this possibility, the difficulty of the learning task was increased in Experiment 2 by reversing the cue-weightings during training for half of the participants (the change groups). For the remainder (the same groups), the cue-weightings did not change. In order to achieve accurate profit predictions, the change groups must engage in learning at least twice: once initially, and once when the cue weights reverse. Thus, this design may be more sensitive to the influence of feedback on learning.

This weighting reversal manipulation may increase sensitivity in a second way, as it changes the task environment from static to dynamic. Full feedback may be particularly valuable in a dynamic environment because it may allow participants to detect any changes earlier, and thus to begin learning about the new environment sooner. Indeed, it has been shown that when the task-structure is abruptly changed, it is very difficult for participants to respond, even if they are given full feedback (Bröder & Schiffer, 2006; Rieskamp, 2006). Thus, Experiment 2 should be sensitive to any benefit offered by comprehensive feedback over selective feedback. It is predicted that the decrease in prediction accuracy due to the cue-weighting reversal will be briefer in duration, and perhaps smaller in magnitude, for the full feedback group than the selective feedback group. Further, the full feedback group is predicted to show more accurate knowledge of the new (changed) cue-weightings than the selective feedback group.

Method

The method for Experiment 2 is similar to Experiment 1, except where noted.

Participants. Forty-eight psychology students from UNSW participated for course credit.

Design. Two between subject manipulations were conducted. The feedback manipulation was the same as in Experiment 1. Orthogonal to this manipulation, the cue-weightings were reversed after the first training block for half of the participants (the change groups) but were not for the remainder (the same groups). Participants were evenly and randomly allocated into these four groups.

Procedure & Materials. The procedure & materials differed from those of Experiment 1 in five respects. First, the base rate of company profitability was 50% for all participants. Second, training was divided into three blocks of 88 trials (rather than four blocks of 60). Third, after the first training block, the weightings of the four cues were reversed for the change groups. Thereby the most heavily weighted cue became the least weighted cue, and the second heaviest weighted cue became the second least heavily weighted, etc. Fourth, the pre-test instructions were altered to explicitly state that no feedback would be given in the test phase.

Finally, we required an online measure of participants' cue-weighting and base-rate estimates, such that it could be delivered both before and after the cue-weighting reversal. To this end, 10 data missing (DM) trials were randomly interspersed in the last 30 trials before the cue-weighting reversal, and a further 10 were interspersed in the last 30 trials of the final training block. There were two types of DM trials: type (i), which assessed the perceived cue-weightings and type (ii), which assessed the perceived base-rate. On each type (i) trial, participants were shown one cue with a maximal score (10) but were told that the data regarding the remaining three cues was missing. There were 2 such trials for each cue, and thus there were 8 type (i) data-missing trials per testing occasion. On each type (ii) trial, participants were told that all information regarding that company had been lost. Two type (ii) trials were given per testing occasion. On all DM trials, as on the normal trials, participants were required to generate a profit prediction. No feedback was given on DM trials.

Results & Discussion

We were primarily interested in how participants would respond to the cue-weighting reversal, and thus it was important that participants had learned the initial cue-weightings. To ensure this, all participants who had a mean error-score of 15 or larger in the first training block were excluded from subsequent analyses. This criterion was selected by examining participants' prediction data from Experiment 1. Fifteen participants were removed, leaving between 7 – 9 participants in each group.

Prediction Data. The accuracy scores for each group are shown in Figure 3. Averaged across all the training blocks, no differences were observed between the same and change groups, $F(1,29) = 2.02$. Within the same groups, no differences in accuracy were seen between those given selective versus full feedback, $F < 1$, replicating Experiment 1. However, within the change groups, the predictions of the selective feedback group were significantly more accurate than those of the full feedback group, $F(1,29) = 7.12$.

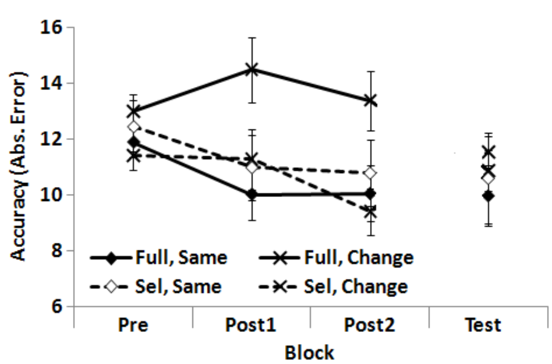

Figure 3: Mean prediction accuracy in Experiment 2.

Averaged across all groups, predictions were more accurate on the last two training blocks (after the weighting-reversal; the Post blocks) than on the training block before the reversal (the Pre block). As expected, the increase in accuracy between the Pre and Post blocks was significantly larger for the same groups (1.71 points) than for the change groups (0.06 points), $F(1,29) = 5.23$. The magnitude of improvement between Pre and Post did not differ between the same groups, $F < 1$. However, amongst the change groups, the predictions of the full feedback group was significantly impaired by the reversal, but the predictions of the selective feedback group were relatively unaffected, $F(1,29) = 4.12$.

Finally, no between-group differences in prediction performance were seen at test, all *F*s < 1. In summary, contrary to our predictions, the cue-weighting reversal did not appear to reduce the prediction accuracy of the selective feedback group. Instead, the reversal specifically impaired the performance of the full feedback group.

Knowledge of task structure. Participants' predictions on the DM trials are summarized in Figure 4. If participants were sensitive to the weightings of the cues, they ought to have predicted more profit on trials in which information was available for the most heavily weighted cue (40%), and least on trials in which information was only available on the least weighted cue (10%). As expected, averaged across groups, a significant linear trend was observed in participants' prediction on the Pre DM trials in which high profits tended to be predicted on trials in which the more heavily weighted cues were shown, $F(1,29) = 41.93$. This trend did not interact with any main-effect, between-group contrasts, all *F*s < 1.This indicates that the groups did not differ in their knowledge of the cue-weights before the reversal.

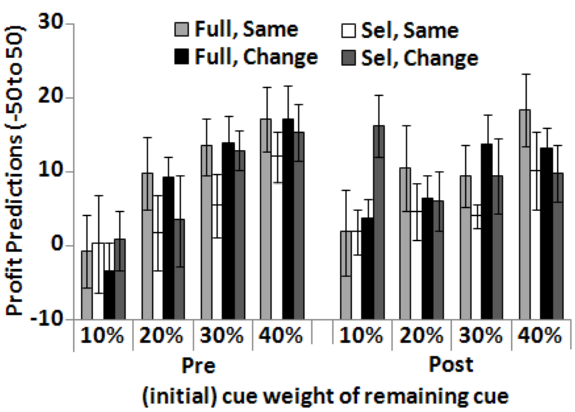

Figure 4: Mean profit predictions on data missing trials in Experiment 2.

The Post DM trials (delivered after the weighting reversal) show a different pattern. Note that for the change groups, that the cue labeled 10% in Figure 4 was the most heavily weighted cue in the Post training blocks. Thus, ideally the change groups should show a negative linear trend on the Post DM trials, while the same group should show a positive linear trend. Averaged across all groups, a significant positive linear trend was observed on the Post DM trials, $F(1,29) = 7.52$. Surprisingly, this linear trend did not interact with the contrast comparing the same and change groups, $F(1,29) = 2.53$. However, it did interact with the contrast comparing the two change groups, $F(1,29) = 4.95$. This interaction reflects the full feedback group's continued use of the initial cue-weighting scheme to produce predictions on the Post DM trials, as contrasted with the selective feedback group's use of the new cue-weighting scheme. Thus, contrary to our hypothesis (but consistent with the training data), the change group given selective feedback learned more about the new cue-weightings than the change group given full feedback. Put simply, full feedback impaired learning in a dynamic environment.

General Discussion

The present experiments examined the benefits of comprehensive feedback in a multiple cue, deterministic prediction task. Consistent with Elwin et al. (2007), Experiment 1 showed that full feedback did not lead to improved prediction accuracy at test, even when the number of training trials was held constant. Further, full feedback did not improve prediction accuracy in training, or increase participants' learning rate (if anything, the opposite was found), or provide better knowledge of the task structure. Experiment 2 showed that full feedback can impair prediction accuracy in training, due to decreased sensitivity to change.

These results are counter-intuitive, and are not consistent with standard error-based models of learning. As noted earlier, these accounts suggest that learning can only occur when feedback is provided. On these accounts, all else being equal, more trials with feedback (or "reinforced" trials) should produce more learning. The present findings go further than prior studies that have shown learning in the absence of prediction error (e.g. Bott, Hoffman & Murphy, 2007) because the present data demonstrate that prediction error can impair learning performance.

One might argue that learning had reached asymptote in the present experiments, and thus further feedback ought not to improve prediction accuracy. However, performance did not approach ceiling in either experiment, as the task was deterministic, and thus perfect performance was possible (in principle). Further, performance does not appear to have been limited by cognitive resources; our computer simulations (not reported here) show that cognitively undemanding strategies could achieve much better performance than was observed. Thus, optimal or asymptotic performance was not reached.

Instead, a more parsimonious account of the data is that, in Experiment 2 at least, selective feedback facilitated learning. This raises a question as to what participants learned on the feedback absent trials. Elwin et al. and Henriksson et al. argue (on the basis of participants' base-rate estimates) that participants simply assume that their prediction was correct on feedback absent trials (termed constructivist coding). Consistent with their hypothesis, base-rate estimates were systematically lower for the selective feedback groups in the present task. However, if participants simply assume that they are correct unless shown otherwise, this should lead to systematic decreases in prediction accuracy, especially in an environment where the cue-weightings were abruptly reversed. No such impairment was observed.

However, a combination of this idea with two further assumptions may account for the present data. First, following Elwin et al. let it be assumed that (i) participants believe themselves to be correct on feedback absent trials and consequently do not encode the details of these trials into working memory. Further, suppose (ii) that participants can only hold information about a small number of trials (a focal set) in working memory at any time (Miller, 1956). Finally, suppose (iii) that participants are more easily able to reason about trials in which high cue-values are shown than on trials in which low cue-values are shown. This assumption may be analogous to the greater ease with which participants learn on cue-present than on cue-absent trials (which has been suggested to explain the differences between cue-competition and retrospective revaluation; Van Hamme & Wasserman, 1994).

If one accepts these assumptions, then the selective feedback condition may facilitate learning by maximizing the efficiency of limited working memory resources. Specifically, selective feedback allows participants to populate their working memory solely with examples which are easy to reason about (high cue-value instances), and this may lead to more rapid learning of the cue-weights and hence more accurate profit predictions. Perhaps the most straightforward test of this speculative account is to modify Experiment 2 such that the selective feedback group only gets feedback on trials in which they predict a loss. According to assumption (iii) this should result in impaired learning relative to the full feedback group.

Alternatively, participants in the full feedback groups may have sought to learn about two outcomes, profit and loss, whereas those in the selective feedback group may have focused on only one outcome, profit. In this case, the two outcomes are symmetrical, and thus learning about both is redundant. Nevertheless, participants did not necessarily know this, and those in the full feedback groups may have sought to learn about both. If so, then the greater difficulty of the dual-outcome prediction task may explain why those in the full feedback condition were most affected by the cue-weighting reversal in Experiment 2. A simple test of this account is to recode participants' predictions and the correct answers on to a scale varying from 0% to 100% profit, rather than from 50% loss to 50% profit. Under these circumstances, this account predicts that the difficulty of the task is equated for the selective and full feedback groups, and thus any performance differences in response to change should disappear.

Acknowledgements

We wish to thank Carissa Bonner for collecting the data, and Andy Wills, Chris Mitchell, Michael Waldmann and York Hagmayer for helpful comments on the project. This research was supported by a Discovery Project grant from the Australian Research Council (DP0877510) awarded to the second author.

References

Bott, L., Hoffman, A. B. & Murphy, G. L. (2007) Blocking in category learning. *Journal of Experimental Psychology: General*, 136, 685–699.

Bröder, A. & Schiffer, S. (2006) Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32, 904–918.

Elwin, E., Juslin, P., Olsson, H. & Enkvist, T. (2007) Constructivist coding: Learning from selective feedback. *Psychological Science*, 18, 105-110.

Henriksson, M. P., Elwin, E. & Juslin, P. (2008) What is coded into memory in the absence of feedback? *Unpublished manuscript, Uppsala University, Sweden.*

Hogarth, R.M. (2005). Is confidence in decisions related to feedback? Evidence from random samples of real-world behavior. In K.Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 456–484). New York: Cambridge University Press.

Kruschke, J. K. (1992) ALCOVE: An exemplar-based connectionist model of category-learning. *Psychological Review*, 99, 22-44.

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

Mackintosh, N. J. (1975) A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82(4), 276 – 298.

Nosofsky, R. M. (1986) Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, 115, 39-57.

Oaksford,M. & Chater, N. (2008), *Bayesian Rationality,* Oxford: Oxford University Press.

Rieskamp, J. (2006) Perspective of probabilistic inferences: Reinforcement learning and an adaptive network compared. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32, 1355 – 1370.

Rescorla, R.A. & Wagner, A.R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp.64-99). New York: Appleton-Century-Crofts.

Rumelhart, D. E., McClelland, J. L., and the PDP research group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume I.* Cambridge, MA: MIT Press.

Van Hamme, L. J. & Wasserman, E. A. (1994) Cue competition in causality judgments: The role of nonrepresentation of compound stimulus elements. *Learning & Motivation*, 25, 127 – 151.