

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Attentional Allocation During Feedback: Eyetracking Adventures on the Other Side of the Response

Permalink

<https://escholarship.org/uc/item/66r3x539>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 30(30)

ISSN

1069-7977

Authors

Watson, Marcus R.
Blair, Mark R.

Publication Date

2008

Peer reviewed

Attentional Allocation During Feedback: Eyetracking Adventures on the Other Side of the Response

Marcus R. Watson (marcusw@sfu.ca)

Department of Philosophy,
Simon Fraser University - 8888 University Drive
Burnaby, B.C. V5A 1S6. Canada

Mark R. Blair (mblair@sfu.ca)

Cognitive Science Program & Department of Psychology,
Simon Fraser University - 8888 University Drive
Burnaby, B.C. V5A 1S6. Canada

Abstract

We used an eye-tracker to investigate the allocation of overt attention during feedback on a categorization task. Results suggest several conclusions: (1) Participants spend a significant amount of time attending to stimuli re-presented during feedback, indicating that this re-presentation may play an important role in learning. (2) Participants spend more time attending to the stimulus during feedback on incorrect trials than on correct trials, indicating that incorrect trials may be more important than correct ones. (3) Attentional allocation to the re-presented stimulus during incorrect trials is predictive of subjects' ability to learn the task, even as few as ten trials into the experiment. This study shows that eyetracking studies of feedback are a promising new method of investigating learning processes.

Keywords: feedback; error processing; learning; categorization; eye-tracking

Introduction

When learning new skills it often helps to be told what you are doing right and wrong. Some kinds of learning are simply impossible without feedback (e.g., Ashby & O'Brien, 2007), and in other cases unsupervised learning can be slower and less successful than learning with feedback (Wulf, McConel, Gärtner & Schwarz, 2002). Understanding how and why feedback works, then, is crucial if we want to understand learning in general. This goal has broad importance for educational theory, and also for the development of efficient experimental paradigms where more participants learn the assigned task.

Feedback has been studied in the context of sports and motor skill learning (e.g., Wulf et al., 2002), human resource management (e.g., Herold & Fedor, 2003), educational instruction (e.g., Bangert-Drowns, Kulik, Kulik, & Morgan, 1991), language learning (e.g., Pashler, Cepeda, Wixted, & Rohrer, 2005), categorization (e.g., Ashby et al., 2002), and neuroscience (e.g., Holroyd & Krigolson, 2007; Seger, in press). A number of studies have compared the effectiveness of different kinds of feedback. Because of the wide variety of learning situations being investigated in these studies it is not clear that their conclusions apply across the board, but it is nevertheless worth briefly considering some results.

In motor learning, feedback that draws attention to the external results of a movement is more beneficial than

feedback that pertains to aspects of the movement itself (Wulf et al., 2002). Furthermore, it seems that giving sparse feedback every few trials is more effective than giving feedback every trial. Wulf et al. suggest that this is because overuse of external feedback causes subjects to become dependent on it and to ignore the internal feedback from their own body.

In many paradigms, however, internal feedback is not particularly useful, at least not initially, because participants have no idea what constitutes satisfactory performance. In these situations, external feedback on the accuracy of responses is often crucial to learning. There have been a number of neurophysiological studies on the role of feedback in such cases, which typically focus on where and how error signals influence neural connectivity. One important recent finding is that neural components associated with feedback may be coding differences between outcomes and expectations, rather than errors per se (Holroyd & Krigolson, 2007; Oliveira, McDonald, & Goodman, 2007). Studies also include efforts to localize error processing to specific brain regions for specific tasks (e.g., Seger, 2007).

Error signals and violations of expectancy are clearly important, but there are other aspects to feedback. On information integration categorization tasks, for example, learning is hampered when feedback is not given on correct trials (Ashby & O'Brien, 2007). Positive reinforcement, then, can also be an important part of learning.

Another source of information during feedback is the stimulus itself. In many real-life situations and experimental paradigms stimuli are visible during feedback, and evidence shows that participants are doing something with the stimuli during this time. Bourne, Guy, Dodd and Justesen (1965) showed that participants in a categorization task who were shown stimuli during feedback made fewer errors than participants who were not. Furthermore, when the length of the feedback phase was increased to 29 seconds (!) those participants who were shown the stimulus during this time continued to improve their performance, whereas the performance of other participants was actually impaired. Bourne et al. suggest that these effects are due to participants forgetting fewer aspects of a stimulus if it remains in view during feedback, which prevents them from accidentally forming hypotheses that are incompatible with this stimulus.

Halff (1975) found a similar effect to Bourne et al. (1965), and further found similar effects of stimulus presentation during feedback in a categorization experiment. He also investigated whether these effects it was correlated with participants' accuracies on each trial the trials they were receiving feedback for. Participants who viewed the stimulus during feedback only on correct trials did not learn faster than those who did not view the stimulus during feedback at all (unlike in Ashby & O'Brien 2007). Participants who viewed the stimulus during feedback on incorrect trials only, as well as those who viewed the stimulus during every feedback phase regardless of accuracy, had significantly higher learning rates, but these rates were not significantly different from each other. Halff concludes that participants test different hypotheses during learning, that these are only modified after error, and that re-presenting the stimulus during feedback aids in preserving memory for the stimulus. Like Bourne et al. (1965), he suggests that better memory for the stimulus allows for improved efficiency of hypothesis revision.

The benefits of stimulus presentation during feedback do not only apply to humans performing categorization tasks. Strength and Zentall (1991) found that pigeons' learning rates also improve when stimuli are re-presented during error feedback on matching-to-sample and oddity-from-sample tasks. Martin and Zentall (2005) confirmed these results and eliminated some other possible confounds. Both studies conclude that pigeons benefit from the chance to review the trial stimuli after errors. Neither investigated whether similar effects result from viewing stimuli during feedback to correct trials.

Bourne et al. (1965), Halff (1975), Strength and Zentall (1991), and Martin and Zentall (2005) are the only works we have been able to find that investigate the role of stimulus re-presentation during feedback. It is somewhat surprising that studies of re-presentation appear to be published at a rate of less than once a decade, given how ubiquitous it is in experimental work. Further investigation is clearly warranted.

The present studies were intended to explore participants' use of feedback in categorization tasks. We used a standard categorization paradigm with consecutive trials of stimulus-response-feedback sequences. Our primary interests were in understanding how participants allocate their attention to different aspects of feedback, how this changes throughout learning, and what this might indicate that they are trying to do with the information they access.

We used an eye-tracker to measure participants' overt attentional allocation, which allowed us to record its development throughout feedback. This is an improvement over the previous practice of inferring attentional allocation based on participants' responses, a method that is indirect and gives no fine-grained temporal information. To our knowledge, eyetrackers have not previously been used to study attention during feedback processing. We use mean total fixation durations to response feedback and to features of the re-presented stimulus as our primary data. We will investigate how these factors change throughout learning and how they change during feedback on correct or incorrect trials.

Experiment 1

Method

Participants. Participants were 20 students at Simon Fraser University who received course credit or pay for their participation. All had normal or corrected-to-normal vision. Two participants were excluded from the analysis for failing to meet the learning criterion (described below).

Stimuli. The stimuli were designed to resemble micro-organisms with three lobules (see Figure 1). Each lobule contained a different type of feature, designed to resemble an organelle. These features were presented in the same location throughout the experiment for each subject. Each feature had two subtly differing possible states, allowing for a total of eight different combinations (see Figure 2). Images were 25.4 cm in diameter, subtending approximately 19.2° of visual angle. Features subtended 1.5°-3.0° and were located 9.2° apart. Variations of the micro-organism's background ensured that subjects would see a unique stimulus on every trial.

Not all features were diagnostic of each category (see Figure 2). Category A was defined solely by a particular value of feature 1, Category B by values of features 1 and 2, and Categories C and D were defined by particular values of all three features. The location of each feature and its mapping onto the category structure was counterbalanced to the extent possible given the number of subjects.

Procedure. The experiment consisted of a series of trials in which a stimulus was displayed, participants identified it as belonging to one of the four categories, and response feedback was given. During piloting we found that participants more often fixated features that were nearest the response buttons. To prevent this, response buttons were not displayed during initial stimulus presentation. Trials began with a fixation cross in the center of a black screen. Subjects clicked on the fixation cross and the stimulus

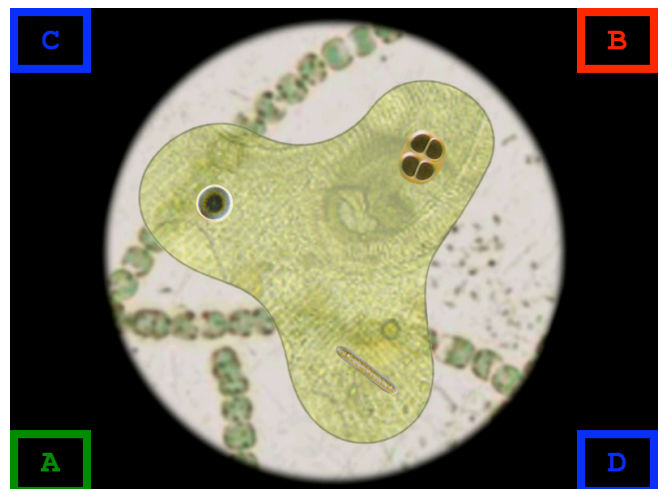


Figure 1: A picture of the feedback screen. This stimulus is an 'A', but the participant incorrectly responded 'B'.

appeared. Subjects viewed the stimulus until they were ready to respond and then clicked the mouse button. The stimulus disappeared and the four response buttons appeared in the corners of the screen, randomly ordered. Once the subject responded, the stimulus was re-presented along with response feedback. On incorrect trials the button that the subject pressed turned red, and on all trials the correct response button turned green. Participants were free to view feedback as long as they wished, and clicked the mouse button to move on to the next trial.

The number of trials varied across subjects. If participants reached a learning criterion of 24 consecutive errorless trials then the experiment continued for a further 72 trials. If they did not reach this criterion by 200 trials the experiment ended.

A Tobii X50 eye tracker sampling at 50hz was used to record gaze data. Fixations were identified using a modified dispersion threshold algorithm and thresholds of 1° and 75ms fixation duration (Salvucci & Goldberg, 2000).

Results and Discussion

We segregated the fixations to the feedback buttons and to the stimulus features. We also split the pre-criterion trials in half, reasoning that the beginning of the session may include strategies that would change once participants became more familiar with the task. Note that this means that the number of trials in the analysis varies widely between participants. We conducted a 2x2x2 analysis of variance (ANOVA) using Stage (first half, last half), Trial Accuracy (correct, incorrect) and Fixation Location (buttons, stimulus features) as within-subjects variables and mean total fixation duration as the dependent measure. Because this measure was non-normal, we rank transformed the data before the analysis.

Results show main effect of accuracy, $F(1,15) = 110.548$, $p < .001$, $\eta_p^2 = .881$, $MSE = 256.196$; Location, $F(1,15) = 467.680$, $p < .001$, $\eta_p^2 = .969$, $MSE = 178.404$ and Stage, $F(1,15) = 87.639$, $p < .001$, $\eta_p^2 = .854$, $MSE = 109.241$; an Accuracy x Stage interaction, $F(1,15) = 11.820$, $p < .01$, $\eta_p^2 = .441$, $MSE = 98.991$ all qualified by a significant 3-way interaction, $F(1,15) = 4.812$, $p < .05$, $\eta_p^2 = .243$, $MSE = 68.224$ a significant three-way interaction, $F(1,17) = 4.81$, $p < .05$, $\eta_p^2 = .24$, $MSE = 68.22$. Results are similar to those shown for Experiment 2 in Figure 3. Overall, participants

spend less time inspecting feedback in Stage 2 of learning than in Stage 1. Participants also spent more time on incorrect trials than on correct trials. This accords with the common sense notion that people do not need to spend as much time processing stimuli that they already know how to categorize. Finally, participants spend more time fixating on stimulus features than they do on response feedback. The importance of stimulus processing is suggested by other research (e.g., Bourne et al., 1965; Half, 1975), and our results confirm this: during feedback before criterion, participants spend about 7 times as long fixating on stimulus features as they do response buttons.

The three way interaction may be easiest to explain as an effect in the processing of stimulus features. When participants go from early learning (Stage 1) to late learning (Stage 2), there is a larger drop in the amount of time spent processing stimuli on correct trials than on incorrect trials. This effect reflects the different roles that correct and incorrect trials play in learning. Early on, when the participant is still learning the relevant stimulus dimensions and the values they can take, re-inspecting the stimuli may be important on all trials, not just incorrect ones. As the participant gets the feel of the task the focus may shift to error-reduction and hypothesis testing and thus to an emphasis on stimulus reprocessing during incorrect trials. While the average time spent fixating to response buttons during feedback does drop throughout the experiment, the drop is roughly equal for both correct and incorrect trials.

Experiment 2

We wanted to replicate the results of Experiment 1 and investigate several new issues to extend our findings. We ran a similar study, but altered the category structure (see Figure 2) to confirm that our results were not due to some peculiarity of the structure used in Experiment 1. We also wanted to make the category structure more difficult in order to produce a greater number of non-learners, which would allow us to investigate differences between learners and non-learners. We reasoned that the new structure would be more difficult because each category is defined by a combination of two feature values, unlike in Experiment 1 where Category A was defined by feature 1 alone (see Figure 2). Finally, we removed feedback after the learning criterion was reached to investigate self-supervised error correction.

Method

Participants. Participants were 32 students at Simon Fraser University who received course credit or pay for their participation. All had normal or corrected-to-normal vision.

Stimuli & Procedure

The experiment was identical to Experiment 1 in all but two respects. First, Categories A1 and A2 were defined by particular values of features 1 and 2, categories B1 and B2 by values of features 1 and 3 (see Figure 2). Second, if the learning criterion of 24 consecutive errorless trials was reached, no feedback about response accuracy was given on subsequent trials. The stimulus reappeared after response

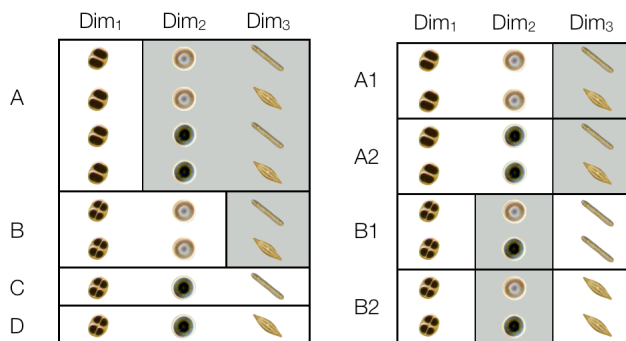


Figure 2: Category structures for Experiment 1 (left) and Experiment 2 (right). Features in gray are irrelevant.

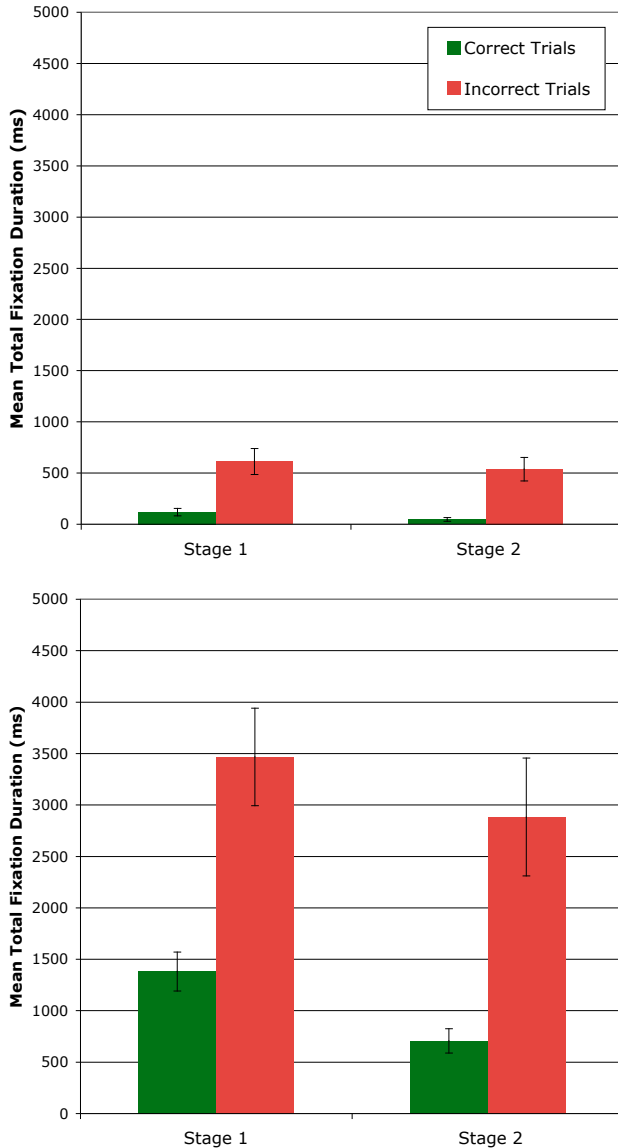


Figure 3: Mean total fixation duration of learners to the feedback buttons (top panel), and the stimulus features (bottom panel) from Experiment 2. Error bars represent standard error.

selection, but the only feedback consisted of whatever button the subject had pressed turning pink. Subjects were explicitly informed that this did not relate to their accuracy.

Results and Discussion

The first analysis was a replication of the analysis of Experiment 1. Using only the data from the 18 participants who met the learning criterion of 24 consecutive error-free trials, we conducted a 2x2x2 ANOVA with Stage (first half, last half), Trial Accuracy (correct, incorrect) and Fixation Location (buttons, stimulus features) as within-subjects variables, again using log total fixation duration during the feedback phase of each trial as the dependent variable.

Results revealed that most of the results from Experiment 1 were replicated with the new category structure in Experiment 2 (see Figure 3). Significant effects included main effects of Accuracy, $F(1,19) = 118.13, p < .001, \eta_p^2 = .86, MSE = 577.55$; Location, $F(1,19) = 307.73, p < .001, \eta_p^2 = .94, MSE = 555.11$; Stage, $F(1,19) = 611.02, p < .01, \eta_p^2 = .367, MSE = 716.81$; the Accuracy x Stage interaction, $F(1,19) = 8.93, p < .01, \eta_p^2 = .43, MSE = 216.28$; the Location x Stage interaction, $F(1,19) = 14.13, p < .01, \eta_p^2 = .426, MSE = 103.64$; all qualified by a three-way interaction, $F(1,19) = 8.77, p < .01, \eta_p^2 = .32, MSE = 48.16$. As in Experiment 1 the three-way interaction seems to stem from the steeper drop in time spent processing stimulus features on correct trials relative to incorrect trials as participants moved from early learning to later learning.

Another of the motivations behind Experiment 2 was to investigate differences between participants who were successful in mastering the categories and those who were not. Increasing the difficulty of the categories had the desired effect: there were 18 learners and 14 non-learners. All non-learners had completed exactly 200 trials, but learners completed 72 trials more than it took them to reach criterion, which varied across subjects. Any comparison across groups such as these involves a measure of arbitrariness, but we reasoned that the two groups would be in the most qualitatively similar circumstances at the beginning of learning.

We conducted a 2x2 mixed ANOVA on data from the first ten trials only, using Learner (learner, non-learner) as the between-subjects variable, Trial Accuracy (incorrect, correct) as the within-subjects variable and mean total fixation time to stimulus features as the dependent measure. There were two subjects who did not make any correct responses in the first ten trials, so we did a least-squares regression substitution for this missing data. A mixed ANOVA with the substituted data revealed a significant accuracy x learner interaction effect, $F(1,34) = 4.46, p < .05, MSE = 1423507.97, \eta_p^2 = .12$.

As shown in Figure 4, though learners and non-learners spend the same amount of time on correct trials, non-learners spend only about 65% of the time that learners spend inspecting the stimuli on incorrect trials, even in the first 10 trials of the experiment. From the very start of the experiment, then, learners were spending much more time processing the stimuli during feedback on incorrect trials than during feedback on correct trials, whereas non-learners were not. This effect is robust enough to allow reasonably accurate prediction of learning success from the first ten trials alone, based on nothing more than fixations to stimulus features during feedback to incorrect trials. 70% of learners have mean fixation durations of over 1990 ms, while 60% of non-learners are below this mark.

Finally, we investigated stimulus feature fixations after learners had reached criterion in both experiments. We found that very few errors are made post-criterion (range = 0 to 3), but that the mean time spent fixating on stimulus features on those incorrect trials ($M=997, SD=867$) is significantly longer ($t=2.85, p<.05$) than on correct trials ($M = 267 \text{ ms}, SD = 266$) in Experiment 1. In Experiment 2, this effect was also significant (correct: $M = 165, SD = 54$;

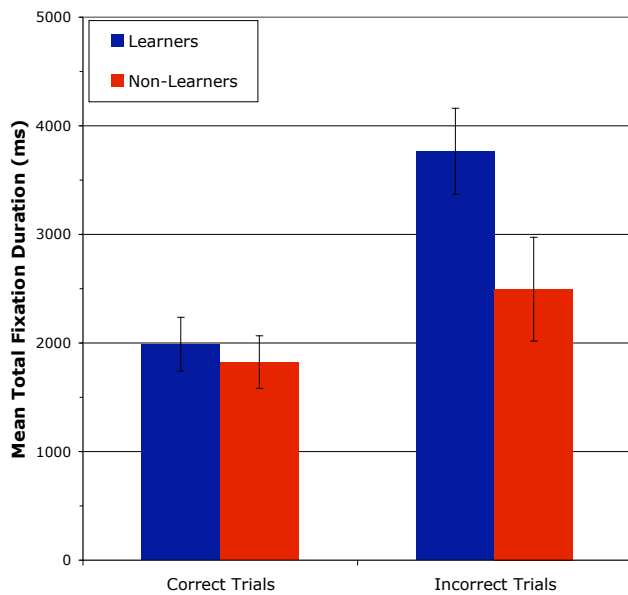


Figure 4: Mean total fixation durations to stimulus features on the first 10 trials of Experiment 2.

incorrect: $M=538$, $SD= 375$; $t=3.127$, $p<.05$). These differences in time spent on incorrect and correct trials in Experiment 2 may be indicative of participants having learned to generate their own error signal, as there is no external feedback post-criterion, and so no way for participants to know they have made an error. This effect will have to be studied more carefully in future research.

General Discussion

The present studies take some first steps toward understanding what participants do when they are processing feedback in a categorization task. The use of eyetracking allowed us to record aspects of stimulus processing that have never before been investigated. There were several findings of note.

First, during the feedback phase of our categorization trials, participants spend more time looking at the stimulus features than the feedback itself on both correct and incorrect trials in all stages of the experiment. This result suggests, in accordance with studies showing that this representation is beneficial to learning (e.g. Bourne et al. 1965; Halfff,1975), that subjects find the re-presented stimulus important.

What is it that further inspecting the stimulus does? There are several non-exclusive plausible possibilities. It may provide a further chance to associate the stimulus with the correct category, since the only time that both the stimulus and the correct category labels are on the screen simultaneously is during the feedback stage. Also, it may provide an opportunity to re-encode the stimulus, either to correct for errors in the initial encoding, or to inspect areas that were not well encoded initially. Alternatively, it may simply prevent the initial encoding of the stimulus from

degrading during feedback. The latter two possibilities are consistent with Bourne et al. (1965) and Halfff's (1975) explanations of their results.

Another interesting point is that the mean total fixation time to the stimulus is larger on incorrect than on correct trials. This ratio differed from early to late learning, such that the relative importance of the stimulus on incorrect vs. correct trials was greater in the later learning trials. This suggests that feedback on incorrect trials is more important to learning than feedback on correct trials, in accordance with most previous work. However, during the initial stage of Experiment 2 learners spend an average of almost two seconds fixating on stimulus features during feedback to correct trials. This suggests that, from the participant's perspective at least, this feedback was far from irrelevant, particularly at the beginning of learning.

Finally, in Experiment 2, as we expected, many subjects do not reach the learning criterion. Learning this task appears to be strongly correlated with the amount of attention paid to the stimulus during feedback on incorrect trials. This effect is present throughout learning, even in as few as ten trials.

The exact reason why some participants do not learn is unclear. Individual differences in learning are, of course, nothing new. Wide differences in learning success have been documented in similar types of categorization tasks (e.g., Blair & Homa, 2001; 2005). More broadly the educational instruction literature identifies differences in feedback propensities (Herold & Fedor, 2003) and motivation/self regulatory skills (Kanfer & Heggstad, 1997) that influence learning. A flurry of recent studies suggest that differences in error processing are identifiable at the neural level. Error processing appears to be impaired in patients with schizophrenia compared to healthy controls in behavioural tasks (Prentice et al., 2007), and patients with schizophrenia also show a decreased amplitude feedback negativity (Morris et al., 2007). Similarly, there is evidence for a reduced amplitude feedback negativity in the elderly compared to younger controls (Eppinger, Kray, Mecklinger & John, 2007). It is unclear at this early stage how any of the neural, behavioural or dispositional variables might relate to our finding of a heightened importance of error trials in learners, but sorting out some of these individual and group differences is obviously an important goal with implications for many fields of research.

How well will our results generalize to other categorization tasks? The categories used in the present studies are complex and differ in several ways from the rule-based and information integration categories that have been previously studied (Ashby & O'Brien, 2007). First, which features are relevant for categorization changes for different categories. This elicits stimulus-dependent attentional shifting from participants and means that stimulus encoding is uneven across stimulus features and categories (Blair, Watson, Maj & Walshe, 2007). It seems likely that such categories encourage the kind of intensive reviewing of stimuli during feedback that we find, perhaps more so than categories with fewer dimensions or categories for which all dimensions are equally important. Another important factor may be the realism of the stimuli used in

the present study. Simpler stimuli and simpler categories may lead to less stimulus re-inspection.

Given the diversity of learning situations involving different types of feedback, both in the laboratory and in the real world, it would be optimistic to believe that our results will generalize to all circumstances. Nevertheless, the various analyses and general approach taken in the present study hold promise for providing a solid set of data that will improve our understanding of how feedback operates in complex cognitive tasks. Further, we would be pleased if these data encouraged more cognitive scientists to concern themselves with what happens on the other side of the response.

Acknowledgments

This work was supported by funding from the Natural Sciences and Engineering Council of Canada, the Canadian Foundation for Innovation, the British Columbia Knowledge Development Fund, and Simon Fraser University to MRB. We would like to acknowledge the Cognitive Science Lab team, specifically: Bill Chen, Fil Maj, Luvdeep Malhi, Kim Meier, Christine Gerson, Gordon Pang, Jordan Shimell, Steve Smith, Calen Walshe, and Mike Wood for aid in data collection, data analysis, and reviewing the relevant literature.

References

- Ashby, F. G., Maddox, W. T., & Bohil, C. J. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory & Cognition, 30*, 666-677.
- Ashby, F. G. & O'Brien, J. B. (2007). The effects of positive versus negative feedback on information-integration category learning. *Perception & Psychophysics, 69*, 865-878.
- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The Instructional Effect of Feedback in Test-like Events. *Review of Educational Research, 61*, 213-238.
- Blair, M. & Homa, D. (2001). Expanding the search for a linear separability constraint on category learning. *Memory & Cognition, 29*, 1153-1164.
- Blair, M. & Homa, D. L. (2005). Integrating novel dimensions to eliminate category exceptions: When more is less. *Journal of Experimental Psychology - Learning Memory and Cognition, 31*, 258-271.
- Blair, M., Watson, M., Maj, F., & Walshe, C. (2007). Extremely Selective Attention: Eye-tracking Studies on Dynamic Attentional Allocation to Stimulus Features. *Psychonomic Society Annual Meeting*, Long Beach, CA.
- Bourne, L. E., Guy, D. E., Dodd, D. H., & Justesen, D. R. (1965). Concept-Identification - The effects of varying length and informational components of the intertrial interval. *Journal of Experimental Psychology, 69*, 624-629.
- Eppinger, B., Kray, J., Mecklinger, A., & John, O. (2007). Age differences in task switching and response monitoring: Evidence from ERPs. *Biological Psychology, 75*, 52-67.
- Half, H. M. (1975). Stimulus presentation after successes and errors in concept-identification. *American Journal of Psychology, 88*, 421-430.
- Herold, D. M. & Fedor, D.B. (2003) Individual differences in feedback propensities and training performance. *Human Resource Management Review 13*, 675 – 689.
- Holroyd, C. B. & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology, 44*, 913-917.
- Kanfer, R. & Heggestad, E. D. (1997). Motivational traits and skills: A person-centered approach to work motivation. *Research in Organizational Behaviour, 19*, 1-56.
- Martin, T. I. & Zentall, T. R. (2005). Post-choice information processing by pigeons. *Animal Cognition, 8*, 273-278.
- Morris, S. E., Mann, M. C., Gold, J. M., & Holroyd, C. B. (2007). ERPs following reward and nonreward in schizophrenia outpatients. *Schizophrenia Bulletin, 33*, 533-533.
- Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior Cingulate is not all error related: Expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience, 19*, 1994-2004.
- Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words. *Journal of Experimental Psychology - Learning Memory and Cognition, 31*, 3-8.
- Prentice, K. J., Robinson, B. M., & Gold, J. M. (2007). Error processing in schizophrenia: The role of WM load. *Schizophrenia Bulletin, 33*, 537-537.
- Salvucci, D.D. & Goldberg, J.H. (2000). Identifying fixations and saccades in eye-tracking protocols. In: Proceedings of the Eye Tracking Research and Applications Symposium 2000,71-78. NY: ACM Press.
- Seger, C. (in press). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neuroscience and Biobehavioral Reviews*.
- Strength, V. & Zentall, T. R. (1991). Matching and oddity learning in pigeons - Effects of penalty time for incorrect responding. *Animal Learning & Behavior, 19*, 49-57.
- Wulf, G., McConnel, N., Gärtner, M., & Schwarz, A. (2002). Enhancing the Learning of Sport Skills Through External-Focus Feedback. *Journal of Motor Behavior, 34*, 171-182.