The dual systems approach to category learning: How do people switch between systems?

Megan Heffernan (mheffernan@psy.unsw.edu.au)
School of Psychology, University of New South Wales
Sydney, NSW, 2052, Australia

Ben R. Newell (ben.newell@unsw.edu.au)
School of Psychology, University of New South Wales
Sydney, NSW, 2052, Australia

Abstract

Dual systems accounts have dominated research into human category learning. Despite this, there has been limited investigation into how people switch between proposed systems. We report an experiment which modified a popular probabilistic categorisation task (the Weather Prediction Task) in an attempt to demonstrate conditions under which such a switch could occur. The results suggested that increasing working memory load impaired categorisation performance and reduced the flexibility with which participants applied their knowledge. Importantly, no clear evidence of a switch between learning systems was found despite using a design intended to favour a shift from declarative to procedural learning. These findings pose questions for the expectation that the nature of a category structure will determine which system is engaged.

Keywords: categorisation; implicit learning; explicit learning

Categorisation is an important cognitive skill which helps people process information about a range of stimuli in day to day life. As such, how people learn to categorise has been the focus of much interest. Research in this area falls into two broad groups: those who argue for a single underlying process (e.g., Lagnado, Newell, Kahan & Shanks, 2006; Newell, Lagnado & Shanks, 2007; Nosofsky & Zaki, 1998; Speekenbrink, Channon, & Shanks, 2008) and those who argue for multiple (typically two) processes (e.g., Ashby & Maddox, 2005; Gluck, Shohamy & Myers, 2002; Knowlton, Mangels & Squire, 1996; Poldrack et al., 2001; Reber, Knowlton, & Squire, 1996). The dual systems approach has dominated research in this area and often rests on the claim that it is the nature of the category structure itself which determines how people will learn a categorisation task (e.g., Ashby & Maddox, 2005).

The dual systems approach argues for a distinction between two systems which underlie category learning. A procedural (or implicit) system which can learn complex, multi-cue category structures but lacks flexibility and is outside of awareness; and a declarative (or explicit) system which can learn simple category structures, produces flexible knowledge and requires awareness. An example of a dual systems model is the Competition Between Verbal and Implicit Systems (COVIS) model (Ashby, Alfonso-Reese, Turken & Waldron, 1998). COVIS has been influential in understanding what might distinguish different category learning systems from each other. Despite the intense research focus on characteristics of the proposed systems, one aspect of the dual systems accounts which remains unclear is how people are able to switch between these systems. The COVIS model claims that there is an initial bias towards engagement of the explicit system, and that the implicit system is engaged only for tasks which are too complex for the explicit system. Although this sets out a general explanation for why a switch might occur between the two systems, the circumstances that predict a switch occurring remain unclear. Understanding *when* and *how* people switch between these systems is crucial to reconciling contradictory findings coming from proponents of single systems accounts and those who favour dual systems accounts. In particular it addresses the question of whether a failure to engage a procedural system is due to inappropriate experimental designs, or the absence of a separate procedural system.

To examine how multiple systems might interact with each other, we will first focus on existing evidence which illustrates the unique characteristics of the proposed systems. One area which provides support for the existence of distinct learning systems has been in probabilistic categorisation, using the *weather prediction task* (WP task). In this task, participants predict a binary outcome (rainy or fine weather) based on four cues. The trial-by-trial learning required to perform the WP task accurately is claimed to recruit the procedural system since information across multiple trials is more useful than information from a single trial (Knowlton, Squire & Gluck, 1994; Knowlton et al., 1996). Consistent with this interpretation is that while participants can learn to correctly predict the outcome they lack flexible knowledge of the cue-outcome associations (Foerde, Knowlton & Poldrack, 2006; Reber et al, 1996).

Poldrack et al. (2001) used the WP task to demonstrate that there are different brain region activations depending on the version of the WP task. The original WP task activated the basal ganglia but left the medial temporal lobes (MTL) relatively deactivated; while a version thought to engage the declarative system had the opposite effect. These findings were interpreted as supporting the idea of competing systems which were engaged based on the nature of the task. This was further supported by Foerde et al's (2006) demonstration of a concurrent task modulating the relative activations of these same areas. This modulation resulted in less flexible cue-outcome knowledge when the MTL was relatively deactivated (concurrent task was present) but if the MTL was activated then participants could demonstrate flexible knowledge. Whether learning

was associated more with striatal or MTL activity had limited impact on accuracy.

While these findings are consistent with a multiple systems account, they also illustrate that young, healthy participants will approach a task designed to engage the procedural system in a way consistent with declarative learning. Such findings could make it difficult to distinguish between 'systems' based on behavioural data because a task designed to engage procedural learning does not always achieve its aim. Thus, findings which illustrate behavioural patterns inconsistent with popular dual systems accounts (e.g., impaired learning under concurrent task conditions Newell et al., 2007) could be explained away by the suggestion that the procedural system was simply not engaged in the first place. How and why this failure to engage might occur are important questions.

Evidence for competition between two systems in the WP task is similar to the COVIS model proposed by Ashby and colleagues. The hypothesised implicit system is reliant on the basal ganglia, while the verbal system relies on the prefrontal cortex. These differences at a neural level are predicted to also be evident in category learning behaviour. Under COVIS, a rule-based (RB) category is one which can be learnt optimally via a simple, verbalisable rule. In contrast, an integration-information (II) category is one where optimal performance requires the accumulation of knowledge about different cues and the application of a rule which is difficult, if not impossible, to report verbally. So, a RB category structure would engage the declarative (or verbal) system and an II structure would engage a procedural (or implicit) system.

The WP task could be considered an II category structure since the optimal rule is difficult to verbalise and it requires the accumulation of information over a number of trials. However, the characterisation of the WP task as an II task is not straightforward due to the different ways the task could be learnt (e.g., Ashby & Maddox, 2005; Gluck, et al., 2002). So although its intention is to be a task which can be best solved using a non-verbalisable rule, it can also be solved to an almost equivalent level of accuracy using a simple rule.

Using the WP task, Foerde et al (2006) illustrated that switching task requirements during learning could elicit results consistent with switching between systems. This suggests that the systems are readily able to decide which one needs to provide the answer on any trial and as such both are possibly in operation at the same time. Consistent with this is Foerde et al's (2006) finding that activity in the striatum remained the same regardless of the presence of the concurrent task. The COVIS model addresses this switching between systems to a certain degree. It claims that there is a bias in favour of the verbal system and that the non-verbal system will be engaged following some level of failure of the verbal system. The competition between the verbal and non-verbal systems is resolved when one delivers a 'stronger' response than the other, and thus that response is used. Although their aim was not to test this exact claim, Poldrack et al's (2001) results using the WP task are consistent with this general idea. They demonstrated greater activation of the MTL initially and then a reduced level of activation as the task proceeded.

Thus, it may be the case that failures to demonstrate procedural learning with the WP task are due to participants failing to 'switch' from a declarative to procedural system. This could explain previous findings that show simple strategies being used to solve the WP task (e.g., Gluck et al., 2002). If these strategies are reinforced to a certain level, participants may lack the necessary motivation to switch to the procedural system. Similar logic has been applied to explain the benefit of early training with difficult items for II categories (Spiering & Ashby, 2008). It was argued that this benefit was due to participants quickly learning that simple RB rules would not correctly categorise items and so they adopted more successful II rules. Conversely, those who were initially trained on easy II items could achieve a sufficient level of accuracy using a sub-optimal rule and so had no need to adopt a more complex II rule. These findings provide another illustration of the effect different category structures have on learning but they do not clearly define the characteristics which necessitate a switch between systems. Past research suggests that the presence of a concurrent task is enough to encourage engagement of the procedural system (e.g. Foerde et al., 2006). Yet the inconsistencies in the literature (e.g. Newell et al., 2007) suggest that this issue remains unresolved.

Most research examining the involvement of separate systems has relied on between-subjects comparisons or the same participants learning different category structures. Our aim was to examine conditions under which healthy, young, adults would switch from a declarative to a procedural system. Would there be evidence for a 'switch' if a simple rule-based approach was initially successful but then led to poor accuracy as the task progressed? To further favour the engagement of a procedural system, a concurrent task was added as the experiment progressed to increase the load on the declarative system. It was hypothesised that if a procedural system was engaged during the task, accuracy would not be impaired even as working memory load and task difficulty increased. Flexibility of knowledge would be poor regardless of the concurrent task if learning was driven by a procedural system. In contrast, if a declarative system is engaged then some cue outcome knowledge should be acquired. However, the presence of a concurrent task will cause a detriment in performance on these measures.

Experiment

Participants

Thirty-eight (30 female) students from the University of New South Wales participated in the experiment in return for course credit. The mean age was 19.9 (*SD* = 3.3).

Stimuli

The current categorisation task was based on the widely used weather prediction task (WP) (e.g., Knowlton et al., 1994; Newell et al., 2007; Reber et al., 1996). The WP task

was adapted to address the criticism that the WP task can be solved by simple rules which inhibits its ability to engage a procedural system (e.g., Ashby & Waldron, 2005; Gluck et al., 2002). To discourage participants from simply memorising patterns which they may have been able to do in the original WP task (4 cues, 14 unique patterns) the number of cues were increased. There were 12 individual cues arranged in sets of three (4 sets total) with 158 unique patterns (see Figure 1). The sets and cues were associated with the outcomes probabilistically (see Table 1). Each cue was assigned a value from -3 to 3 to reflect these probabilities. The outcomes (rain or fine weather) were determined using an information integration rule used in previous research (Price, 2005; Waldron & Ashby, 2001): *If S1(value) + S2(value) + S3(value) + S4(value) > 0 then: Outcome = fine, otherwise, outcome = rain.* For example, if cue 9 (Set 3) had an assigned value of 1, and cue 12 (Set 4) had an assigned value of 1, the outcome for a pattern made up of these two cards would be: *S1(*0*) + S2(*0*) + S3(*1*) + S4(*1*)* = 2. Since the value is greater than zero, the outcome would be fine weather.

Although the outcomes for each pattern were fixed the relationships between cues and outcomes were probabilistic (see Table 1). That is, although a participant might learn that C1 was unlikely to give an outcome of rain, there were patterns where C1 was present *and* the outcome was rainy weather. To maintain consistency with previous WP tasks, patterns consisted of between 1 to 3 cues, all cues were relevant and participants never saw cues from the same sets on the same trial. Patterns which resulted in an outcome of zero were excluded. The sets of cues were weakly associated with the outcomes compared to WP tasks which used 4 cues only (e.g. Newell et al., 2007; Reber et al., 1996).

The category structure was arranged so that there were *set inconsistent* (SI) and *set consistent* (SC) cues. SI cues were those with an outcome-association inconsistent with the overall set outcome association. For example, from Table 1 it can be seen that Set 4 is the most strongly associated with the rainy outcome (P(rain) = .65). Within this set however, only cues C10 and C11 are strongly associated with rain (P(Rain) = .87). Cue C12 is inconsistent with this because it is strongly associated with the fine outcome (P(Rain) = .22). Such an arrangement increases the importance of focusing on the multiple individual cues rather than applying a simple set-level rule, and thus should favour the engagement of a procedural system.

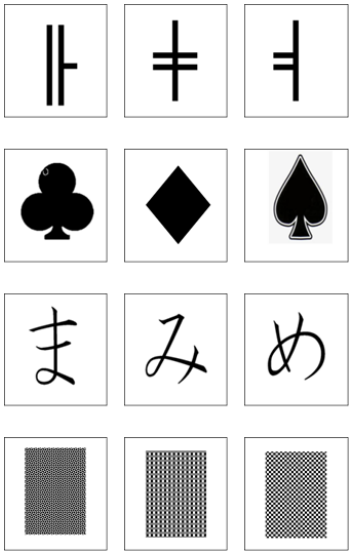

Figure 1: Stimuli used for categorisation task. Each row represents a set of cues.

Table 1: Category structure, showing set probabilities, cue probabilities and frequencies.

Set	Set P(rain)	Cue Type	Frequency	P(rain)
S1	.35			
		C1	45	0.13
		C2	45	0.13
		C3	50	0.78
S2	.45			
		C4	53	0.31
		C5	53	0.31
		C6	49	0.76
S3	.55			
		C7	53	0.69
		C8	53	0.69
		C9	49	0.24
S4	.65			
		C10	45	0.87
		C11	45	0.87
		C12	50	0.22

Design and Procedure

Weather Prediction Task A two groups design was used. A Load group completed the WP task as well as a concurrent task. A No Load group completed the WP task only. The WP task consisted of 40 trials which involved patterns made up of SI cues (10 unique patterns, 4 cues only). This initial blocks of trials was aimed at encouraging the use of simple rules. Then 234 trials were shown which were made up of patterns using both SI and SC cues (158 unique patterns, including the 10 already shown). In these trials, participants would not be successful at predicting the outcome if they tried to apply what they had learnt about the SI cues to similar looking SC cues. Rather, if there is a procedural system tracking cue-outcomes it should start to make the predictions because it would be more accurate. Trials were shown in a random order that was fixed across participants. There were 274 trials in total. The stimuli used are shown in Figure 1 and were randomly assigned to the different set values.

Participants were told that they would be predicting the weather based on a set of tarot cards. They were told that although they might need to guess at the start, they would

eventually become better at predicting what the weather would be based on the presented tarot cards. Participants in both groups were also told that they might also be asked to perform a numbers task while they were performing the weather prediction task.

Numerical Load Task Participants in the Load group completed a concurrent task at the same time as the WP task. The concurrent task has been used previously to impair the declarative system (e.g., Newell et al., 2007; Waldron & Ashby, 2001). In this task, participants see two numbers between 1 and 9 on either side of the WP task cues. These are on-screen for 1 sec and one number is physically larger than the other. On each trial participants are asked to report which number was *physically* larger or which number had the *higher value.* Participants see these numbers at the same time as the cue presentations. The Load group performed the concurrent task on trials 31 to 274. The concurrent task was started towards the end of the SI patterns rather than at trial 41 (when SC cues were added) to ensure participants had little reason to associate the presence of the concurrent task with the appearance of additional patterns and cues.

Flexibility of Knowledge Participants were given two measures of flexibility of knowledge (Reber et al., 1996). The *cue estimation* task required participants to estimate the likelihood of an outcome given a particular cue. Participants saw all three cues from each set at the same time and were asked to rate on a sliding scale what the likely outcome was. The markers on the scale were "Extremely likely to be fine", "As likely rainy as fine" and "Extremely likely to be rainy". The order of the presentation of the sets was randomised across participants. In the *cue selection* measure participants were shown 4 individual cues (1 from each set) and asked *"If you knew it was going to be fine (or rainy) and only one card was showing, which is the most likely to be showing?"* Participants were asked this question for both outcomes (counterbalanced) and for SI and SC cards (order randomised).

Results

Numerical Load Task Performance of 80% accuracy on the numerical load task was applied as exclusion criteria (consistent with Waldron & Ashby, 2001). Five participants from the Load group were excluded and their data is not included in subsequent analyses (Load n = 17). There was no exclusion applied to the No Load group (n = 16). Mean accuracy in the numerical load task was 90.1% (*SD* = 4.2) which was significantly above chance ($t(16) = 39.14$, $p < .05$).

Overall Learning Overall learning performance (Figure 2) shows that both groups performed similarly and to quite a high level during the initial two blocks. There was a drop in performance at Block 3 which is consistent with the additional cues being added to the task. Numerically the No Load group were performing better than the Load group in blocks 3-11, although both groups appeared to be improving across training. A 2(condition) x (11)(block) repeated measures ANOVA confirmed that there were significant

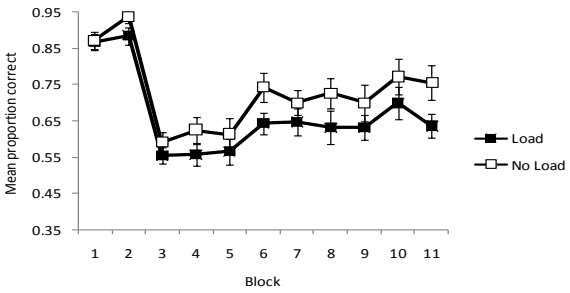

Figure 2: Mean WP task performance for both groups across blocks (error bars = 1 SEM). Blocks 1 and 2 = 20 trials each and consist of *set inconsistent* patterns only. Blocks 3-11 = 26 trials each and consist of *set inconsistent* and *set consistent* patterns.

linear and quadratic trends (ps < .05). These reflect the expected learning patterns. However, there was no main effect of condition ($F(1,31) = 3.2$, $p = .084$) which is inconsistent with the prediction that the Load group would be impaired by the presence of the concurrent task. No other effects were significant.

The lack of group difference may be due to both groups having equal opportunity to learn about the SI patterns during the first 30 trials (concurrent task started at trial 31). This may have helped the Load group for subsequent trials involving SI cues. To investigate this possibility, SI and SC trials were analysed separately. SI trials (Figure 3, panel A), are those where the predicted outcome using set probability was different to the outcome predicted using the probabilities of individual cues. SC trials are those where the group level and individual cue predictions are the same (Figure 3, panel B). Separate repeated measure 2 x (9) ANOVAs were carried out on the SI trials and SC trials for blocks 3-11. For SI trials there was a significant quadratic trend ($F(1,31) = 4.3$, $p < .05$). The main effect of group and interactions were not significant (Fs < 1). These results are unsurprising given the mostly equivalent learning conditions in the first two blocks of the WP task. For the SC trials, there was a significant linear trend ($F(1, 31) = 50.11$ $p < .05$) and a main effect of group ($F(1,31) = 4.85$, $p < .05$). No other results were significant. Results for the SC trials suggest that increased load on the declarative system did impair the Load group's ability to learn about cues which had not featured in the initial phase of the task. This is consistent with a single system account but is inconsistent with the engagement of a procedural system.

Flexibility of Knowledge In the cue estimation measure accuracy of estimates was assessed by firstly calculating a regression slope for each participant based on their estimates for each cue. This slope was compared to the slope of an 'ideal' participant whose cue estimates were the same as the objective estimates (slope = 12.11). If participants' regression slopes did not differ significantly from this it would be taken as evidence for good accuracy on the cue estimate measure. Secondly, slopes for the No Load (M = 11.64, *SD* = 7.18) and Load (M = 1.00, *SD* =

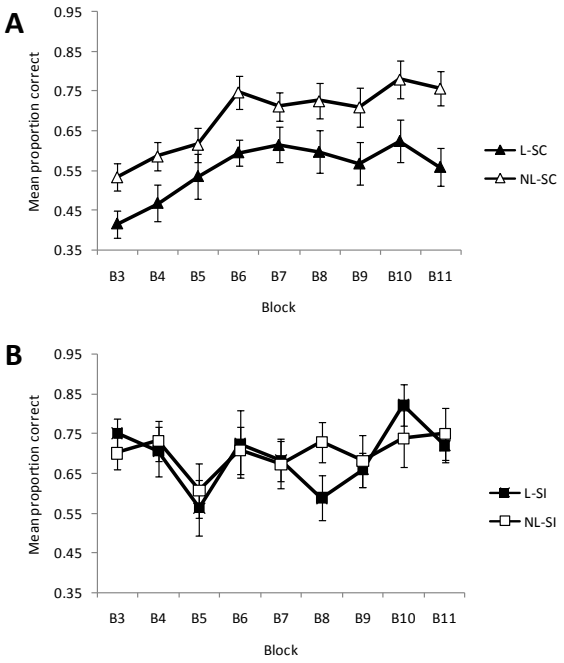

Figure 3: Mean WP task performance for both groups on set inconsistent (Panel A) and set consistent (Panel B) trials. Only data for blocks 3-11 are shown because there were no SC trials in Blocks 1-2 (error bars = 1 SEM).

5.43) groups were compared to assess whether the concurrent task impaired performance. Calculated slopes for the Load group were significantly different from the ideal participant ($t(16) = 3.50$, $p < .05$) while those for the No Load group were not ($t(15) = .34$, $p > .05$). This indicates that participants in the No Load group were more accurate in their estimates. This impairment for the Load group is consistent with procedural learning since they did not achieve a basic level of cue estimation accuracy. Further to this, the mean slopes for both groups were significantly different from each other ($t(31) = 2.53$, $p < .05$) illustrating that the concurrent task impaired the Load groups' knowledge of cue-outcome associations. These two results indicate that the Load group may have engaged a procedural system *and* that there was an additional impairment caused by the concurrent task. The No Load group showed performance consistent with declarative learning.

Cue Selection data was scored using the method described in Reber et al. (1996). Scores ranged from 1 (best) to 4 (worst). The mean cue selection score for SI cues was 1.69 (SD = .73) and 1.83 (SD = .86) for the No Load and Load groups respectively. The mean cue selection score for SC cues was 1.78 (SD = .98) and 1.88 (SD = 1.05) for the No Load and Load groups respectively. These scores were all better than chance (ts > 2.46) which is consistent with declarative learning. If the concurrent task caused a detriment in declarative learning then it would be expected that the No Load group would perform better on SC cues compared to the Load group. Although the No Load group were numerically better, a 2(Load vs. No Load) x (2)(SC vs. SI) ANOVA showed that there were no main effects (Fs < 1.11) and no significant interaction ($F < 1$).

Discussion

This experiment used a modified version of the weather prediction task to test for a switch from a declarative to a procedural system. Results indicate that while participants under additional working memory load were not impaired overall, their ability to learn about additional cues as task difficulty increased was impaired. Measures of cue knowledge support the conclusion that the No Load group were reliant on a declarative system to learn this task. However, results were mixed for the Load group. The cue estimation measure provided the best evidence for the engagement of a procedural system, yet this was not supported by the cue selection measure which demonstrated that the Load group had acquired better than chance cue-outcome knowledge. The contradictory results obtained from two measures which are expected to index the same type of knowledge needs to be further investigated to better understand what is driving this difference. Potentially this reflects the difficulties in getting accurate measures of explicit knowledge (see e.g. Lagnado et al., 2006). Alternatively, it may be indicative of a difference in sensitivity within measures of cue knowledge which need to be addressed to ascertain whether it is truly reflective of a lack of declarative knowledge.

The lack of difference in overall learning accuracy suggests that the impact of the working memory task was limited. This is consistent with the dual systems account and potentially indicative of a switch to procedural learning. However, the modified version of the WP task allowed us to examine these differences more closely. There was clearly no difference between the two groups in patterns where predictions ran counter to the predicted set relationship (SI trials). This is not surprising given that the two groups were completing *only* the WP task during initial trials made up of SI cues. In contrast, when the task increased in difficulty the Load group was impaired for patterns that were not shown during the first two blocks. It was expected that increased task difficulty combined with a simple rule that was no longer effective, *and* increased working memory load, would lead to learning consistent with the engagement of a procedural system. However, this impairment (compared to No Load) suggests that the Load group were continuing to learn the WP task via a declarative system.

Foerde et al (2006) reported that healthy, young adults were inclined to rely on the declarative system to learn the WP task. Our results are consistent with that finding. However, Foerde et al. also illustrated that a concurrent task would result in a switch away from use of the declarative system. Why participants would continue to rely on a system which was failing in light of increased task difficulty and increased working memory load in our task is unclear. Under a dual systems account such as COVIS there should have been a switch to procedural learning. As such, our results are more consistent with single system accounts of category learning (e.g., Newell et al., 2007). Despite

conditions which are thought to favour procedural learning and impair the declarative system, participants failed to demonstrate a switch in systems. Findings such as these illustrate that although dual systems accounts dominate interpretations of category learning data there remain important, unanswered questions and inconsistencies within the literature (see also Lagnado et al., 2006; Nosofsky & Zaki, 1998; Speekenbrink et al., 2008).

Our overall results provide evidence against the idea that the detriment caused by the concurrent task is simply the result of a failure to switch from declarative to procedural learning. For such an interpretation to hold, there would need to be a clear account for why a procedural system was not engaged in the current categorisation task. Conditions favoured the engagement of a procedural system, yet participants continued to rely on their declarative system. These results are consistent with single system accounts which do not posit a unique role for procedural learning. In contrast, for the current data to fit with a multiple systems account a conception of category learning that places individual characteristics as a key determiner of which system is engaged would need to be developed. For example, that in certain individuals (e.g., healthy adults) the declarative system is so dominant or favoured that it needs to be seriously compromised before the switch to a procedural system will be made. If that is the case, there needs to be further attention paid to identifying which characteristics are the most crucial. For example, whether accuracy has to be below a certain point or if individual differences are influencing the switching between systems. This interpretation does not reflect current multiple systems accounts. At present, dual systems accounts argue that the relative engagement of these systems depends on the category structure (e.g. Ashby et al., 1998). Such a claim does not satisfactorily account for the pattern of data found in the current experiment.

One area for future research is examining whether imposing a time-limit for responding would have an effect in the modified WP task. Past research (e.g., Foerde et al., 2006) used time limits but the current experiment did not. Time limits are thought to favour the procedural system because the declarative system does not have the time to engage in the necessary hypothesis-testing (Ashby et al., 1998). So although the current experiment aimed to favour procedural system engagement through increased task difficulty and increased memory load, this is a potential shortcoming and worthy of future investigation.

In conclusion, a modified version of the WP task illustrated that even in conditions favouring procedural learning participants do not switch away from a declarative system. There remain unanswered questions about the interaction between systems in dual systems accounts. These need to be addressed to more completely understand what underlies category learning.

References

Ashby, F.G., & Maddox, W.T. (2005). Human Category Learning. *Annual Review of Psychology, 56*, 149-178.

Ashby, F.G., Alfonso-Reese, L.A., Turken, A.U., & Waldron, E.M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105*, 442–481.

Gluck, M., Shohamy, D., & Myers, C. (2002). How do people solve the "Weather Prediction" task? Individual variability in strategies for probabilistic category learning. *Learning and Memory, 9*, 408 – 418.

Knowlton, B.J., Squire, L.R. & Gluck, M.A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory, 1*, 106-120.

Knowlton, B.J. Mangels, J.A. & Squire, L.R. (1996). A neostriatal habit learning system in humans. *Science, 273*, 1399 -1402.

Lagnado, D.A., Newell, B.R., Kahan, S., & Shanks, D.R. (2006). Insight and strategy in multiple cue learning. *Journal of Experimental Psychology: General, 135*, 162-183.

Meeter, M., Myers, C.E., Shohamy, D., Hopkins, R.O. & Gluck, M.A. (2006). Strategies in probabilistic categorisation: Results from a new way of analysing performance. *Learning & Memory, 13*, 230-239.

Newell, B.R., Lagnado, D.A. & Shanks, D.R. (2007). Challenging the role of implicit processes in probabilistic category learning. *Psychonomic Bulletin & Review, 14*, 505-511.

Nosofsky, R.M. & Zaki, S.R. (1998). Dissociations between categorisation and recognition in amnesics and normal individuals: An exemplar based interpretation. *Psychological Science, 9*, 247-255.

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature, 414*, 546–550.

Price, A.L. (2005). Cortico-striatal contributions to category learning: Dissociating the verbal and implicit systems. *Behavioural Neuroscience, 119*, 1438-1447.

Reber, P. J., Knowlton, B. J., & Squire, L. R. (1996). Dissociable properties of memory systems: Differences in the flexibility of declarative and nondeclarative knowledge. *Behavioural Neuroscience, 110*, 861-871.

Shohamy, D., Myers. C.E., Onlaor, S. & Gluck, M.A. (2004). Role of the basal ganglia in category learning: How do patients with Parkinson's disease learn? *Behavioural Neuroscience, 118*, 676-686.

Speekenbrink, M., Channon, S. & Shanks, D.R. (2008). Learning strategies in amnesia. *Neuroscience and Biobehavioural Review, 32*, 292-310.

Spiering, B.J. & Ashby, F.G. (2008). Initial training with difficult items facilitates information integration, but not rule-based category learning, *Psychological Science, 19*, 1169-1177.

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning. *Psychonomic Bulletin & Review, 8*, 168–176.