National and Kapodistrian University of Athens
Department of Philosophy and History of Science
Interdisciplinary Program of Graduate Studies in Basic and Applied Cognitive Science

The Effect of Cue Naming in Probabilistic Category Learning

Fotios Fotiadis
08M20

First Supervisor: Prof. Athanassios Protopapas
Second Supervisor: Dr. Argiro Vatakis

Athens, 2010

First supervisor:

Prof. Athanassios Protopapas

Signature

____

Second supervisor

Dr. Argiro Vatakis

Signature

____

Index

Preface

I would like to express my gratitude to Prof. Athanassios Protopapas for accepting to be my supervisor and for always being there for me: encouraging me, correcting me, urging me to work. I would also like to thank Dr. Argiro Vatakis for being a "second" supervisor only in name and for providing me with her immediate support and guidance on any occasion.

My deepest thanks go to Prof. Martijn Meeter for providing all the details of the WPT experimental set-up and for kindly replying to all of my questions. My gratitude goes also to Prof. Lori Holt (and her student Sung-Joo Lim) for providing the auditory stimuli used in the third experiment of the present study. Many thanks to Prof. M. A. Gluck, Prof D. R. Shanks, Dr. D. Lagnado, and Dr. B. Newell for replying to my e-mails.

Last but not least, I would like to thank Eleni Vlahou for her instrumental assistance with DMDX software and with data analysis.

Abstact

Learning in a prototypical probabilistic category learning task, the Weather Prediction Task, has been thought to be mediated by both the procedural and the declarative memory systems, presumably at different periods in training. It has been suggested that healthy participants approach the task by simple verbalizable rules mediated by the declarative memory systems, at least early in training. Based on neurocognitive and philosophical accounts we hypothesized that an operating characteristic of declarative knowledge is the availability of the contents of knowledge to be verbally expressed. We manipulated the ability of participants to verbally express the names of the auditory cues of the task and hypothesized that participants in a condition utilizing difficult-to-name computer-generated tones (non-nameable condition) would be unable to develop, or rely on, declarative verbalizable rules early in training, as opposed to participants in a condition utilizing easy-to-name animal sounds (nameable condition). We predicted, based on previous reports on healthy participants' performance, that performance of participants in the non-nameable condition would be at chance levels for a longer period, compared to performance of participants in the nameable condition. Indeed such a between-subjects difference in performance was observed and we therefore believe that the present research provides an experimental manipulation that hinders the engagement of the declarative memory system in learning, at least early in the task.

Part 1
Introduction

1.1. Weather Prediction Task

1.1.1 Multiple Memory Systems Hypothesis

Memory is not considered to be a single human faculty. The hypothesis of Multiple Memory Systems (MMS) has been introduced by many neuroscientists (see Squire, 2009, for a historical review) and is considered to be one of the most successful ideas in the field of cognitive neuroscience (Poldrack & Foerde, 2008). Long-term human memory is thought to be comprised by multiple distinct systems and support for this idea, with respect to human cognition, has been originally offered by dissociations in performance observed in memory tasks executed by amnesic patients (Foerde & Poldrack, 2009).

Memory systems are defined as distinct in terms of anatomy but "also in terms of operating characteristics, the kind of information processed and the purpose served by each system" (Squire & Zola, 1996, p. 13515).

Originally a dichotomous view of memory systems was adopted. Binary distinctions were made such as "declarative" vs. "procedural" knowledge or "explicit" vs. "implicit" memory. Given the diversity of phenomena though, researchers shifted away from such a two-part dichotomy and assumed the existence of more than two memory systems. A broader distinction (see Fig. 1.1) has emerged between "declarative" and "non-declarative" memory or learning systems (Squire, 2004).

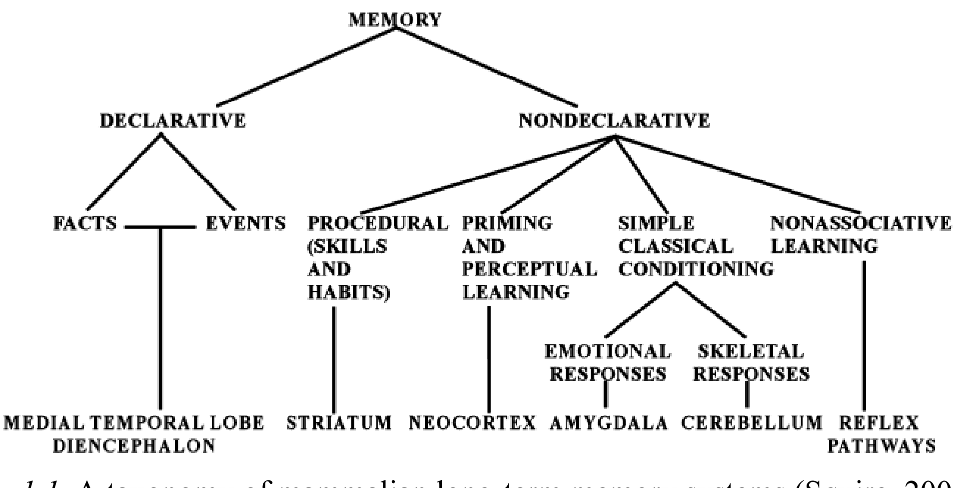

Figure 1.1. A taxonomy of mammalian long-term memory systems (Squire, 2004, p. 173)

One of the tasks that have been used in cognitive neuroscience to assess the contribution of distinct memory systems to the incremental acquisition of cognitive skills is the Weather Prediction Task (WPT; Poldrack & Rodriguez, 2004). Two systems are in the focus of the studies employing the task. The declarative memory system, subserved by the medial temporal lobe (MTL) and related diencephalic structures, which is thought to support memories for facts and events. Also, the procedural memory system, subserved by the basal ganglia, which is thought to support the acquisition of motor and cognitive skills and habits (Squire, 2004).

1.1.2. Early neuropsychological studies

The WPT was introduced in 1994 in a neuropsychological study with amnesic patients. (Knowlton, Squire, & Gluck, 1994). The task is an adaptation of a task previously used to study probabilistic classification learning (Gluck & Bower, 1988). It is a categorization task in which participants classify stimuli appearing on a computer screen in one of two possible categories.

Four cues comprise the stimuli in this task. Each cue is independently associated to one of two outcomes with a fixed probability and the outcomes appear equally often throughout the task. The cues are combined with each other so that a fixed number of possible combinations ("patterns") is produced. On each trial one of these combinations is presented to the participant and he/she is asked to predict the outcome. Finally, participants receive feedback on their response (Knowlton et al., 1994).

The task was thought to discourage the use of a declarative learning strategy because of its probabilistic structure. The participant cannot depend on memorizing the association between a combination of cues and an outcome on a specific trial since the actual outcomes may vary across trials. Thus participants have to gradually accumulate information of the cue-outcome associations across trials based on feedback. The task was therefore considered by Knowlton et al. (1994) to "provide an analogue in human subjects for the habit learning tasks by animals" (p. 106) and to be, therefore, procedurally mediated.

In Knowlton et al. (1994) three tasks were administered to a group of amnesic patients and to three control groups. Different cues and number of trials were used in each task (see Fig. 1.2).

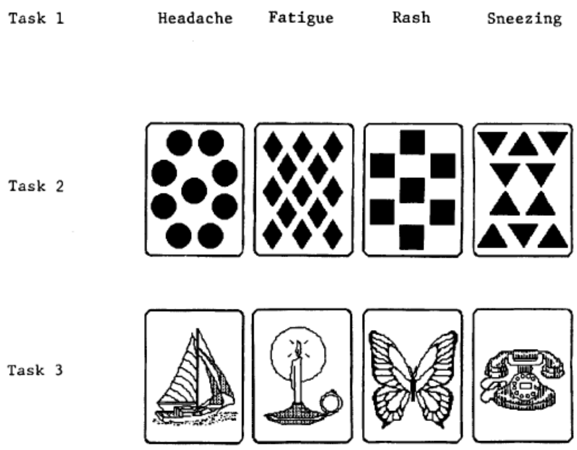

Figure 1.2. The four cues used for each of the three tasks in the study by Knowlton et al., 1994 (p. 110).

In Task 1 amnesic patients and control participants saw one, two, three, or four symptoms on each trial and they had to decide whether an imaginary patient exhibiting these symptoms would have one of two fictitious diseases. The task consisted of 350 trials. In Task 2 the cues were cards with geometric shapes (rounds, squares, diamonds, and triangles) and the two possible outcomes were “sun” and “rain”. This task became the prototypical task that is now called the “Weather Prediction Task”, and in the Knowlton et al. (1994) study consisted of 350 trials. The same outcomes, “sun” and “rain” were also used in Task 3, but the cues were images of objects (boat, candle, butterfly, telephone device). In this third task the number of trials was 50.

After the categorization task both groups were administered a questionnaire that was thought to assess declarative knowledge of the task. This test task consisted of 11 multiple-choice questions about the nature of the cues, the layout of the computer screen, and the training episode (Knowlton et al., 1994).

Performance of the participants was measured in terms of optimal responding. The participant was considered to have made the correct choice if he/she had chosen on each trial the outcome that was, throughout the task, most often associated with the cue pattern presented. Performance was analysed in blocks of 10 trials for the first 50 trials (considered to reflect early learning) and in blocks of 50 trials for the remaining trials (considered to reflect later learning). Chance performance is 50% correct since the two outcomes occurred equally often, and responses on cue patterns for which both outcomes were equally likely were excluded from analysis.

The data revealed that both amnesic patients and control individuals performed similarly in the 50 first trials on each task. Their performance was near chance in the first 10 trials and was significantly above chance for trials 41 to 50. A three-way ANOVA on data pooled from the three tasks revealed a significant main effect of trial block but no effects of either group or task and no interactions were found. For Tasks 1 and 2, in which training was extended past 50 trials, differences emerged in performance between the amnesic and control group. Although in the final block of 50 trials both groups scored significantly above chance, the performance of the control group was better than that of the amnesic patients.

As for the questionnaire assessing declarative knowledge of the tasks, performance of the amnesic patients group was impaired relative to performance of the control group for all three tasks.

The analysis of the data showed that in early learning both groups performed similarly, although declarative knowledge of the task was worse for the amnesic group than for the control group. Thus, Knowlton et al. (1994) were able to show a single dissociation and concluded that “declarative knowledge does not contribute to the early acquisition of classification learning” (p. 114). The fact that the control group outperformed the amnesic patients later in training was attributed to the acquisition of declarative knowledge of the task in later trials by the control group. Overall, Knowlton et al. (1994) concluded that the contribution of the hippocampal system resulting in the acquisition of declarative knowledge could only become apparent in later training, whereas in early training participants acquired nondeclarative knowledge of the cue-outcome relationships.

The result of similar performance between an amnesic and a control group in early learning during the WPT was replicated in a subsequent seminal study (Knowlton, Mangels, & Squire, 1996). A 150-trial version of WPT was administered to two patient groups, an amnesic patients and a Parkinson's Disease (PD) patients group, and to a control group. After training, all groups were administered a questionnaire of eight multiple-choice items thought to assess declarative memory for the classification task.

The researchers tested for a double dissociation between declarative memory and habit memory and indeed the experimental data provided evidence in favour of their hypothesis.

Analysis of behavioural data for the first 50 trials showed that amnesic patients and control participants gradually learned the cue-outcome associations and that both groups' performance improved from near chance in the first 10 blocks to above chance in the 5th block of 10 trials. A two-way ANOVA of the data of the two groups revealed a significant effect of block but no effect of group or interaction. On the other hand, the PD patients exhibited no significant improvement in performance across 50 trials. The PD group performance was no better than chance in the 5th block of 50 trials. A two-way ANOVA on the data of the control and of the PD group revealed an effect of group on performance. Moreover, when comparing the data of a subset of the PD patients—the ones with the most severe symptoms, called PD* in the study—and the data of the amnesic patients by means of a mixed ANOVA, a significant interaction of group x block was revealed.

Performance on the test task showed that amnesic patients were unable to acquire declarative knowledge of the task whereas PD and PD* patients exhibited normal declarative memory for the classification task, as compared to the control group.

These results were considered by the researchers as providing evidence for a double dissociation of memory functions between the structures damaged in amnesia and the structures damaged in PD. They concluded that "probabilistic classification learning depends on the neostriatum but not on the medial temporal lobe or diencephalon, and the opposite is the case for declarative memory" (Knowlton et al., 1996. p. 1400). These conclusions could be drawn only for the first 50 trials, since further training showed that both PD patients and control participants improved their performance across the final 100 trials. Knowlton et al. (1996) attributed this improvement in performance to declarative knowledge that participants might have been able to acquire with further training.

The researchers stated that the neostriatum is a brain structure that is important not only for motor behaviour and learning but also for the gradual incremental learning of associations. They considered it to constitute a parallel and separate learning system with respect to other brain systems such as the MTL, the amygdala, the cerebellum, and the neocortex (Knowlton et al., 1996).

1.1.3. Neuroimaging

Significant advance in elucidating the contribution of the basal ganglia and the MTL in probabilistic category learning was made by a functional magnetic resonance imaging (fMRI) study in 2001 (Poldrack et al., 2001).

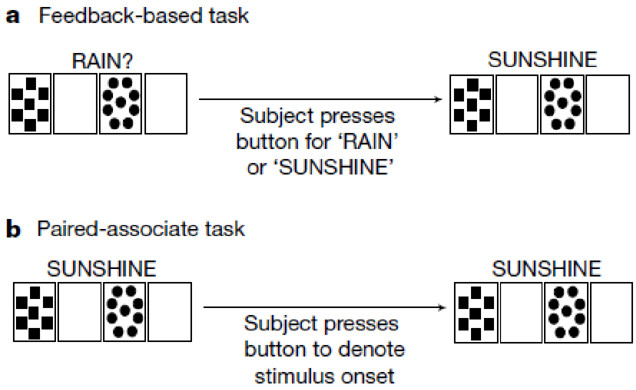

Figure 1.3. Depiction of the Feedback-based and the Paired-associate tasks in Experiment 1 (Poldrack et al., 2001, p. 547).

Poldrack et al. (2001) conducted two experiments. In Experiment 1, fMRI was used to image brain activity of healthy participants that took part in one of two versions of the WPT (see Fig. 1.3). The first version, called "feedback-based" (FB), was the typical WPT in which participants are assumed to learn the cue-outcome associations through corrective feedback and is thought to be mediated by the procedural memory system (as clearly stated by Knowlton et al., 1996). A second version, called "paired-associate" (PA), was designed to emphasize declarative memory processes. In this version the outcome was presented at the same time as the stimulus and remained on the screen throughout the trial. Participants were requested to respond by pressing a key, simply to denote the appearance of the stimulus, and thus to equate for any perceptual and motor processes present in the typical task. This second version is thought to engage declarative memory (Poldrack et al., 2001). Performance of the participants in the second task was assessed immediately after the last scan by presenting each pattern twice and asking for a category label; no feedback was provided. A baseline task was also included in the design, with similar perceptual and motor characteristics, but no learning demands.

Behavioural results showed that performance was similar in both versions suggesting that despite the differences in the experimental set-up the knowledge acquired in both versions was capable of supporting comparable accuracy. Analysis of the fMRI signals in the FB versus the baseline task revealed a significant activation of the basal ganglia and a deactivation of the MTL (see also Poldrack & Foerde, 2008, for a meta-analysis of neuroimaging data from four studies confirming the aforementioned pattern in brain activity during the task). Further analysis compared brain activity during the FB task with the one observed during the PA task and the results showed that MTL exhibited increased activity in PA task whereas the caudate nucleus exhibited increased activity in the FB task. Moreover, correlation analysis for each participant's BOLD signal change showed that activity in right caudate nucleus was found to be negatively correlated with the activity of left MTL, providing evidence for a negative relationship between these two structures.

In order to test this—possibly competitive—interaction between the two brain structures, another group of 14 healthy subjects executed 96 trials in the FB task using event-related fMRI over the course of two scans (Experiment 2). Examination of parametric changes in the evoked response over time showed that “MTL was initially active and caudate was initially inactive, but that the MTL quickly became deactivated and the caudate nucleus became activated” (Poldrack et al., 2001, p. 547).

These results were taken by the researchers as providing evidence for competition of the two memory systems in the neuronal level. Besides this important conclusion, the neuroimaging study by Poldrack et al. (2001) provided a new proposal as to the engagement of the declarative memory system in probabilistic category learning. For the first time it was suggested that “the MTL is engaged very early in learning and that it then becomes deactivated throughout training” (p. 548). The researchers acknowledged the procedural differences in the tasks used in their study and in the early neuropsychological studies suggesting the opposite pattern of engagement for the declarative memory system, but they claimed that their results are consistent with both animal studies and neurocomputational models of category learning.

In particular, according to a computational theory of category learning (Gluck & Myers, 1993), the hippocampal module develops new representations of the cue-cue regularities. Once these representations are formed, they then can be used by other modules such as the neocortex or the striatum. Thus, the initial activation of the MTL and its subsequent deactivation, found by Poldrack et al. (2001), was consistent with the predictions made by the Gluck, Oliver, & Myers (1996) cortico-hippocampal model of category learning (Poldrack et al.).

1.1.4. Strategy analysis

The notion that learning in WPT is mainly mediated by nondeclarative memory systems was challenged by a strategy analysis (Gluck, Shohamy, & Myers, 2002).

Gluck et al. (2002) claimed that performance levels of various clinical groups in the task do not elucidate the neural substrates that support learning. It has been assumed in the past that learning of the cue-outcome associations is acquired incrementally, "as if there were four independent conditioning processes going on in parallel" (p. 409). Yet, the performance levels typically reported by control groups (70% to 75% on average) in previous studies (Knowlton et al., 1994,1996) could be achieved through a simple strategy focusing on the memorization of one (of the total four) cue-outcome associations. Gluck et al. presented techniques, on the basis of post-hoc analyses of behavioural responding, that might reveal qualitative differences in the ways in which different populations approach the task.

Gluck et al. (2002) applied these post-hoc analyses to healthy participants' performance on a replication of the WPT (Experiment 1) and on a newer version of the task, which was thought to maintain the probabilistic structure of the task while making the cue-outcome association more discriminable (Experiment 2). This "newer" version is the one most often used since (e.g., Hopkins, Myers, Shohamy, Grossman, & Gluck, 2004; Shohamy, Myers, Onlaor, & Gluck, 2004, but see Foerde, Knowlton, & Poldrack, 2006; Foerde, Poldrack, & Knowlton, 2007; Price, 2009) and it is thought to reduce participants' frustration (Hopkins et al., 2004). The cue-outcome associations used in the newer version are depicted in Fig. 1.4.

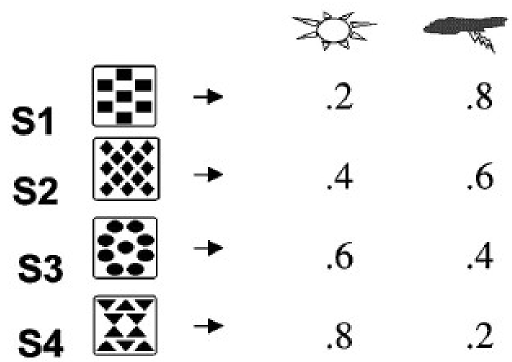

Figure 1.4. Cue-outcome association used in the Gluck et al. (2002) "newer" version of WPT (Shohamy et al., 2004, p. 678).

In Experiment 1, participants were administered a questionnaire immediately after the categorization task. The questionnaire included an open-ended question in which participants were asked to describe the strategy they used in the task. A multiple-choice question in which they had to select the answer best describing their strategy was also included and, finally, two questions assessing the participants' estimates of the cue-outcome associations were administered.

The participants' performance across 200 trials of the older version of WPT was 62.41% and in the newer version it was 74.5%. This increase in performance in the newer version was predicted and expected since the new cue-outcome associations, according to Gluck et al. (2002), made the task easier to learn (but see Foerde et al., 2006, for an alternative interpretation). A repeated-measures ANOVA on both tasks revealed a main effect of block on performance and no effect of gender or block by gender interaction.

Based on participants' responses to the open-ended question in Experiment 1, Gluck et al. (2002) posited four strategies (see Fig. 1.5) that were assumed to be used by participants in approaching the task. They constructed for each strategy an ideal participant's performance and checked which strategy best fitted each participant's performance on every block of 50 trials.

Table 4. *The Rule for Constructing Ideal Subject Data for Each Strategy Investigated*

Strategy	Ideal subject data constructed by:	% Optimal
Multi-Cue	Assume the optimal response (i.e., most frequently-correct outcome) is made on every trial.	100%
Singleton	Assume the subject learns the optimal response for each of the four singleton patterns and guesses on the remaining trials.	75%
One-Cue (highly predictive)	Assume the subject is responding based on the presence or absence of one of the highly-predictive cues (square or triangle) and ignoring all other cards.	87.5%
One-Cue (less predictive)	Assume the subject is responding based on the presence or absence of one of the less-predictive cues (diamond or circle) and ignoring all other cards.	66%

(% Optimal) = potential performance by a subject reliably following this strategy through the entire experiment.

Figure 1.5. Table depicting the four classes of strategies thought to be used more often by participants in the WPT (Gluck et al., 2002, p. 413).

Apart from the multi-cue strategy which allows an ideal participant to score 100% optimal responses, the remaining three strategies defined by Gluck et al. (2002) are sub-optimal, and are thought to depend on rapid, non-incremental learning processes. According to Shohamy, Myers, Kalanithi, & Gluck (2008) reviewing the Gluck et al. findings, these sub-optimal strategies may be thought to be mediated by the declarative memory system.

The results of the strategy analysis showed that in both the older and the newer version of WPT, a sub-optimal (and presumably declarative) strategy provided a better fit than the optimal strategy for 90% of participants (Shohamy et al., 2008). Participants apparently started by using a simple singleton strategy and gradually shifted towards a more optimal multi-cue strategy.

Yet, the results of the strategy analysis were not correlated with the participants' answers to the questionnaire. For example, some of the participants stated in the open-ended question that they had been using a multi-cue strategy, but their knowledge of the cue-outcome associations was incorrect and their performance (in the last block of 50 trials) suggested that they had been using another, simpler, strategy.

This observation led Gluck et al. (2002) to suggest that even easily verbalizable rules (such as “Respond 'rain' whenever the triangle card is present”) that reflect simple sub-optimal strategies might be “acquired in an unconscious, nonverbalizable way” (p. 416). Thus strategy analysis might not be able to fully distinguish between declarative and non-declarative learning processes.

Gluck et al. (2002) concluded that “a normal healthy human may use a variety of strategies and brain structures—possibly in parallel—to approach a difficult categorization task” (p. 416). Thus, the notion that learning in WPT is mainly supported by nondeclarative memory systems has been questioned, at least when young healthy individuals approach the task.

1.1.5. Neuropsychological Findings Revisited

The neuroimaging study by Poldrack et al. (2001) suggested that the declarative memory system mediates performance during the first trials of the WPT. This notion renders the intact performance of amnesic patients in these first trials (as shown in the early neuropsychological studies by Knowlton et al., 1994, 1996) a puzzling finding that needed further examination.

Indeed, a study with amnesic patients in 2004 led to different findings (Hopkins et al., 2004). The group of patients recruited in the study suffered from amnesia and the etiology was hypoxic injury which is thought to cause fairly selective damage to the hippocampus. All patients underwent magnetic resonance imaging and a bilateral hippocampal damage was verified. Amnesic patients and matched control participants were tested on the WPT and also on a variant of it, called the "ice cream task".

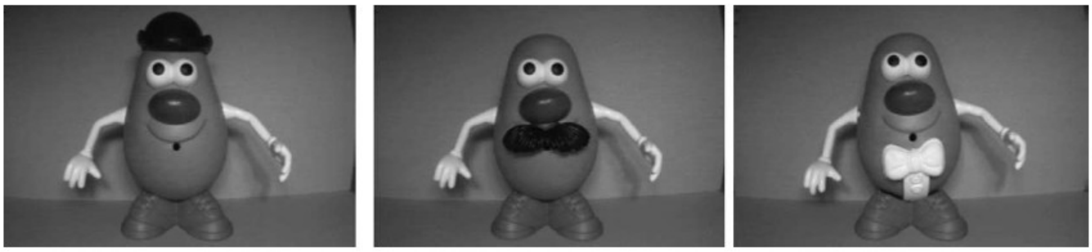

Figure 1.6. Examples of stimuli presented in a variant of the WPT, the "ice cream task" (Hopkins et al., 2004, p. 529).

The ice cream task had the same logical and probabilistic structure as the WPT but uses a different stimulus set (see Fig. 1.6). On each trial of the task an image of Mr. Potatohead appears on the screen. The participant is asked to guess whether Mr. Potatohead, considered to be a customer in an ice cream shop, wants vanilla or chocolate. The prediction is based on four cues, four facial features that may or not be added to the basic figure: a black hat, a black mustache, red eyeglasses, and a white bow tie. The rest of the set-up regarding the feedback to the participant's response is identical to the one in WPT (Hopkins et al., 2004).

Overall, across the 200 trials on both tasks, the amnesic group's performance was worse than the control group's performance. The data obtained for the early stages of the categorization tasks were analyzed in 5 blocks of 10 trials. For the WPT, a repeated measures ANOVA revealed that the effect of group on early performance fell short of statistical significance. Yet, when the data from one control participant (who achieved less than 50% correct in the first 50 trials) were excluded from the analysis, there was a significant effect of group on performance. For the ice cream task, a similar analysis revealed that there was a significant effect of group on performance.

The difference in findings between the Hopkins et al. (2004) and the Knowlton et al. (1994, 1996) studies was attributed to the different etiologies of amnesia. In the Knowlton et al. (1994) study the etiology of amnesia was hypoxic injury, diencephalic damage, and in one patient the cause of amnesia was unknown. Moreover, the control group in the Hopkins et al. study was quite younger (by 20 years on average) compared to the Knowlton et al. control group. Hopkins et al. reported pilot studies showing that older subjects perform worse than younger subjects, and suggested that age might have contributed to the masking of the differences between amnesic and control group in the early neuropsychological studies. Finally, Hopkins et al. implemented the "newer" version of the experiments, following Gluck et al. (2002), which renders the categories in the tasks slightly more discriminable as compared to the Knowlton et al. version.

Hopkins et al. (2004) concluded that "it certainly appears possible that medial temporal damage affects probabilistic category learning, at least under some procedural circumstances, and in some amnesic etiologies" (p. 534).

This same year, the findings of the Knowlton et al. (1996) study regarding the performance of PD patients were also challenged. A strategy analysis (following Gluck et al., 2002) was performed on the behavioural data of a PD patients group (Shohamy et al., 2004).

The researchers tried to clarify the nature of the contribution of the basal ganglia to probabilistic category learning. They hypothesized that basal ganglia damage would impair learning on the WPT due to either a quantitative effect, namely a slower learning rate (in which case the same strategies were expected to be used by both the PD and the control group), or due to qualitative effect on the availability of the two memory systems (resulting in different strategies adopted by the two groups). The 200-trial "newer" version of WPT (Gluck et al., 2002) was administered to a PD and to a control group for three consecutive days.

Analysis of behavioural data showed that both groups significantly improved across the 3 days of training and that PD patients were impaired compared to controls on the second and third day of training. Strategy analysis revealed that control subjects shifted from a simple to a more complex optimal learning strategy, whereas the majority of PD patients continued to rely on simple sub-optimal learning strategies (that focus on single cues) throughout the three days of the experiment. Hence, PD patients failed to learn the optimal learning strategy for the task and the effect of damage of the basal ganglia on probabilistic category learning seemed to be qualitative rather than quantitative (but see Shohamy et al., 2008, for an alternative account).

Although the study focused on strategy analysis, the researchers conducted an analysis of the two groups' performance on the first 50 trials of Day 1. In contrast to the reports of Knowlton et al. (1996), there was no significant difference between the groups in this period. A repeated-measures ANOVA found that the main effect of group on performance fell short of statistical significance. Moreover, as noted earlier, PD patients were impaired relative to control subjects not on the first day of training, but on the second and third day. Thus the impairment observed in this study seemed to be a “late training deficit” (Shohamy et al., 2004, p. 683) as opposed to the early training deficit observed in the Knowlton et al. study.

Again, this contrast was attributed to the profile of the PD group. While in the Knowlton et al. (1996) study some of the PD patients had severe symptoms, in the Shohamy et al. (2004) study the patients were diagnosed with mild to moderate Parkinson's disease. This difference was considered important since the disease affects progressively different brain structures and hence PD seems to have differential effects on cognitive functions depending on its progress (Shohamy et al.).

The same lack of early training deficits in a group of PD patients with mild Parkinsonian symptoms, compared to a control group of healthy individuals, was also reported in a neuroimaging study (Moody, Bookheimer, Vanek, & Knowlton, 2004). In this study participants alternated between WPT blocks and control task blocks in 4 fMRI runs, and completed throughout the experiment 96 trials of the WPT; the probabilistic structure utilized was that of Knowlton et al. (1994).

The main finding of the study was a differential activation of MTL in the two groups. While in control participants a right MTL region was deactivated during the categorization task (a finding also reported by Poldrack et al., 2001), in PD patients the same region of interest was activated. The authors suggested that this finding supports the hypothesis of competitive interaction between the declarative and the procedural memory systems and suggested that the task is approached by PD patients in an explicit (declarative) way since the basal ganglia in this clinical population are compromised.

Behavioural data of the WPT showed that both groups improved in their categorization ability across trial blocks. Yet, as in the Shohamy et al. (2004) study, the PD patients did not exhibit impaired performance compared to healthy individuals. The contrast between the finding of this study and the Knowlton et al. (1996) study was again attributed to the fact that the patients in the former study had earlier stage PD than the patients in the latter study

1.1.6. Secondary-task effects.

The encoding of memories in healthy participants during the WPT by either the declarative or the procedural memory system was studied by Foerde et al. (2006). The authors followed a methodology often used to manipulate the acquisition of implicit knowledge in another (presumably) procedurally mediated task, the Sequence Learning task (Nissen & Bullemer, 1987). A secondary task, thought to occupy working memory, was introduced to the WPT. This distraction task had been assumed to decrease declarative memory encoding while leaving performance levels intact (Foerde et al., 2007).

Participants were trained on two similar classification problems while they were scanned using fMRI. They were trained on one problem under single-task (ST) conditions and on the other under dual-task (DT) conditions, that is, while performing concurrently with the WPT a tone-counting task. The classification tasks were structurally isomorphic, differing only in the color of the geometric shapes of the cards. Learning of the categories was based on trial-by-trial feedback, as typically in the WPT. After training the participants received additional probe trials in which they classified items that had been trained under either ST or DT conditions. No feedback was provided during these probe trials, preventing further learning of the categories. After scanning, the participants' declarative knowledge about the cue-outcome associations was assessed through questionnaires (Foerde et al, 2006).

Analysis of the performance during training revealed only a marginal impairment in the second block of 50 trials for DT conditions. On the probe tests accuracy was similar for the two (ST and DT) conditions. The test of declarative knowledge revealed that the secondary task impaired acquisition of flexible knowledge about cue-outcome associations (Foerde et al. 2006). Thus the behavioural results are in accordance to other studies (Foerde et al, 2007, but see Newell et al., 2007) showing a dissociation between classification performance (which was not affected by DT conditions at encoding) and declarative knowledge (which was affected).

Comparison of the fMRI signals during DT versus ST training conditions revealed additional activity in auditory and prefrontal cortices in the former condition. These results were consistent with the notion that DT conditions were more demanding. Activity in the striatum did not differ between conditions, supporting the notion that "habit-learning mechanisms are automatically and obligatory involved when a task is performed" (Foerde et al., 2006, p. 11780). A significant finding emerged through analysis of the fMRI signals during the probe trials; the analysis revealed significant differences in neural activity between items learned under DT versus ST conditions (see Table 1, Foerde et al., 2006) providing evidence for the hypothesis that the relative contribution of the declarative memory and the habit learning system on PCT may be modulated by performance on a secondary task. Further analyses examined the relation between behavioural performance and brain activity during probe tests. A double dissociation was revealed: activity in the right hippocampus was significantly correlated with performance on items learned under ST but not under DT conditions, whereas the opposite pattern was found for activity in the putamen which was significantly correlated with performance on items learned under DT but not under ST conditions. Similarly, activity in both the right and left MTL was significantly correlated with the measure of declarative knowledge (as assessed by the questionnaires) for items learned under ST conditions but no such correlation was found for DT-items (Foerde et al., 2006).

According to the authors the results demonstrated that equivalent levels of learning can be supported by either the MTL or the striatum. It appeared that—at least for healthy participants—"the PCT can be learned by using different brain systems depending on the task demands" (Foerde et al., 2006, p. 11781).

1.1.7. Contemporary research and open questions.

The WPT continues to be a paradigm widely used in cognitive neuroscience in order to clarify the differential contribution of the declarative and the procedural memory system to the acquisition of cognitive skills (Poldrack & Foerde, 2008).

Some researchers, though, have questioned the implicit nature of the knowledge acquired in WPT (Lagnado, Newell, Kahan, & Shanks, 2006) and challenge the interpretation of the task as a procedurally mediated task (Newell, Lagnado, & Shanks, 2007; Price, 2009). It has been suggested that perhaps MMS accounts of previous findings need to be re-examined (Newell et al., 2007), whereas others have argued in favour of the MMS hypothesis since it seems able to account for human behavioural data whereas alternative theories do not (Poldrack & Foerde, 2008).

Many neuropsychological studies employing the WPT have elaborated on the nature of deficits found in clinical populations (Shohamy, Myers, Grossman, Sage, & Gluck, 2005; Shohamy et al., 2008; Wilkinson, Lagnado, Quallo, & Jahanshahi, 2008). A question that remains open is one regarding the nature of impairment evident in PD patients' performance. Behavioural data from the Shohamy et al. (2004) study indicated different strategies employed by PD patients compared to control subjects. This finding gave rise to qualitative explanations: (1) presumably PD patients have to rely on alternative memory systems in order to learn (Shohamy et al., 2004) or (2) PD impairs an assumed strategy-shifting cognitive mechanism (Shohamy et al., 2008). Simulation, though, of PD patients' performance by Gluck and Myers (1993) cortico-hippocampal model of category learning, by means of a numerically lower learning rate, was capable to replicate the observed experimental behaviour (Shohamy et al., 2008). The modeling data suggested that a quantitative explanation (a slower learning rate) could account for the lack of shift in strategies and overall performance levels of PD patients. Thus, further examination is needed to distinguish between qualitative and quantitative accounts of the impaired performance of PD patients and check the model's prediction that the aforementioned patients might acquire the optimal strategies with much extended training.

A matter of current debate addressed by neuroimaging studies, embracing the MMS hypothesis, is the nature of the interaction between the distinct systems. Some studies suggested a competition in the neuronal level between the MTL and basal ganglia during training (Moody et al., 2004; Poldrack et al., 2001). An alternative interpretation of neuroimaging findings involves the two systems contributing in parallel to learning (Shohamy et al., 2008). Indeed, in Foerde et al. (2006) it was suggested that the declarative and the procedural memory systems may acquire redundant information. The researchers suggested that, according to their findings, a potential competition appears to be occurring “when the knowledge is applied rather than during acquisition of the task” (p 11782). Thus, new studies addressing the nature and timing of a purported competition between distinct memory systems are required.

Moreover, despite numerous experiments employing the WPT, there is no unanimous understanding of how participants solve the task. Behavioural data have been analyzed and simulated by mathematical models in order to provide insight on the issue. Strategy analyses (introduced by Gluck et al., 2002, and refined by Meeter, Myers, Shohamy, Hopkins, & Gluck, 2006) assumed that performance is governed by rule-based cognitive mechanisms. Rolling regression (Lagnado et al., 2006) was based, though, on an alternative assumption, namely a general learning mechanism that integrates multiple cue-outcome contingencies. Both assumptions (rule-based learning and incremental learning) have been proven to be equally successful in predicting performance of both healthy and clinical populations but some evidence emerged in favour of rule-based learning due to the fact that it predicted (as opposed to the assumption of incremental learning mechanisms) the observed non-linear performance observed in some individuals (Meeter, Radics, Myers, Gluck, & Hopkins, 2008). Clearly, further inquiry into the cognitive mechanisms mediating performance in the WPT is needed—a matter that according to Ashby & Mattox (2005) is strictly empirical—, and also, the need for a new conceptualization of probabilistic category learning has emerged (Meeter et al., 2008).

To conclude, according to the interpretation of findings in the WPT by Foerde & Poldrack (2006), two different types of learning (declarative and procedural) may be acquired within the task. While in the early neuropsychological studies (Knowlton et al., 1994, 1996) it was suggested that learning in the WPT was mediated by the procedural memory system, the current understanding differs:

> "in the healthy brain, multiple cognitive processes and multiple neural systems may contribute to learning, and a given task can most likely be learned in more than one way" (Shohamy et al., 2008, p. 222).

The manipulation of the engagement of each system during learning in the WPT has proven to be fruitful in the elucidation of the assumed interaction between the declarative and the procedural memory systems. Experimental manipulations of learning conditions (Poldrack et al., 2001), secondary-task demands (Foerde et al., 2006; 2007), or timing of feedback (Foerde, Clement, & Shohamy, in preparation), have provided evidence in favour of the MMS hypothesis. The present study continues on this effort and introduces an experimental manipulation that is thought to hinder the contribution of declarative memory system, at least early in training. The manipulation in question is based on one of the characteristics of declarative knowledge which is further examined in the following sections.

1.2. Declarative knowledge

1.2.1. On the nature of declarative knowledge

Although the declarative memory system has been on the focus of scientific inquiry since the mid 50's (Scoville & Milner, 1957) and the declarative-procedural distinction is a central one according to proponents of the MMS hypothesis (e.g., Poldrack & Foerde, 2008), researchers do not share a common definition of the long-term memories encoded by the declarative memory system, that is, of declarative knowledge, in terms of its operating characteristics. *Declarative* knowledge has been characterized as *conscious* (e.g., Schacter, 1987), *explicit* (e.g., Reber & Squire, 1994), and *flexible* (e.g., Reber, Knowlton, & Squire, 1996). Although these terms have been used rather interchangeably, clearly their use suggests different conceptualizations on the nature of declarative knowledge. Moreover, sometimes declarative knowledge is defined in terms of its differences to procedural knowledge in a more or less circular manner (since the *procedural* memory system is a *nondeclarative* memory system, e.g., Squire & Zola, 1996).

According to Tulving (1985), who introduced to the study of memory a comparable distinction between procedural, semantic, and episodic memory systems, each assumed memory system, besides being supported by distinct neural structures, “differs in its method of acquisition, representation, and expression of knowledge” (p. 387). An examination of the term declarative as introduced by pioneers in the field of learning and memory is attempted in the following to shed light on these three suggested aspects (acquisition, representation, and expression) of declarative knowledge.

1.2.1.1. What did “declarative” originally mean?

The declarative-procedural distinction was first introduced in artificial intelligence (McCarthy & Hayes, 1969) and was later used in cognitive modeling by Anderson to reflect two qualitatively different faculties of cognition (e.g., Anderson, 1976, as cited in Anderson, 2007). According to the ACT-R theory, cognition arises from the interaction of declarative and procedural knowledge. Declarative knowledge items (*chunks*) are created by the encodings of objects in the environment, whereas procedural knowledge items (*production rules*) are created by the encodings of transformations in the environment (Anderson, 1996).

The same dichotomy was introduced in the study of memory by Cohen & Squire (1980) in a seminal paper examining preserved learning of amnesic patients in a mirror-reading task. According to the authors, declarative knowledge is *data-based* knowledge, whereas procedural knowledge is *rule-based* knowledge, and their findings were considered to support the hypothesis that a related distinction is present in the nervous system.

D. L. Schacter, a pioneer in the study of implicit memory, citing Squire & Cohen (1984), stated that "conscious or explicit recollection is a property of a declarative memory system" (Schacter, 1987, p. 511). Since then, the declarative-procedural dichotomy is often very closely related to the explicit-implicit dichotomy (e.g., Reber & Squire, 1994; Squire, 1992; Squire & Zola, 1996).

The fact that declarative knowledge is thought to be acquired in a data-based manner and that its nature seems to be explicit and conscious, reviewed so far, seems to underlie a common feature: the ability of the holder of the knowledge to verbalize its contents. Such an account of items of declarative knowledge was stressed by Squire (1992). In a paper reviewing evidence from animal studies he stated that (human) "declarative memory seems to imply an ability to declare one's knowledge verbally" (Squire, 1992, p. 204).

In the years to come, the conceptualization of declarative knowledge as explicit complicated the definition of *declarative* further. The confusion arose from the fact that *declarative* refers to an assumed memory system (Squire & Zola, 1996) whereas *explicit* refers to experimental test criteria. According to Reed & Johnson (1994) examining the Sequence Learning task, "implicit learning is demonstrated when indirect tests reflect knowledge acquisition and direct tests do not" (p. 585). Direct tests are those that encourage performance on the basis of consciously available information, whereas indirect tests avoid encouraging the use of conscious knowledge. Adding to the confusion, much debate has emerged as to the appropriateness of test criteria thought to assess implicit learning (for a review see Dienes & Perner, 1999).

Finally, Squire and Zola (1996) reviewed evidence in favour of the assumption that declarative memories are more flexible than procedural memories. Reber et al. (1996), in an experiment employing the WPT, assessed flexibility of knowledge by the fact that it could support performance in novel experimental contexts.

Following Tulving's theorization of memory systems (Tulving, 1985) and the conceptualization of declarative knowledge previously reviewed, we may conclude the following: Declarative knowledge is knowledge supported by the MTL and related diencephalic structures and it is acquired in a data-based manner (Cohen & Squire, 1980). It is thus representational knowledge, in that declarative memories represent something (facts or events; Squire & Zola, 1996) in the real world. The knowledge acquired is explicit or available to conscious recollection (Schacter, 1987), it is flexible in that it can manifest its existence in novel contexts (Reber et al., 1996), and it can be verbally expressed (Squire, 1992).

This last characteristic of declarative knowledge, the availability of items of declarative knowledge for verbal reports, is the focus of the present study. In the following, a philosophical account that seems to strengthen the notion of declarative knowledge as verbally expressible knowledge is reviewed.

1.2.1.2. The explicit-implicit distinction and declarative knowledge.

Dienes & Perner (1999) attempted to clarify the explicit-implicit distinction in terms of the functional properties of mental representations and offered a theory of implicit and explicit learning. Based on the representational theory of mind by Fodor (e.g., Fodor, 1978, as cited in Dienes & Perner, 1999), they considered knowledge to be a propositional attitude, arising from the functional role of the assumed corresponding internal representation. The characterization of knowledge as explicit or implicit may vary according to which of the three aspects of the propositional attitude (self, attitude, and content) is explicitly represented.

Dienes & Perner (1999) linked the explicit-implicit distinction to related distinctions and experimental test criteria, in an effort to provide a unified experimental methodology and relate the findings of studies in research areas such as learning, memory, perception, and cognitive development. When considering declarative knowledge they characterized it as “a close relative of verbally expressible knowledge” (p. 741), and they suggested that declarative knowledge is knowledge that states its predication and factuality explicitly (*content-explicit*) as opposed to procedural knowledge which is contained in the application of a procedure. They also analyzed the verbalization criterion often used to assess explicit knowledge in terms of direct experimental tests. They suggested that knowledge that is communicated verbally represents its contents explicitly too. Thus, according to the "theory of implicit and explicit knowledge" by Dienes & Perner (1999), verbal communication and declarative knowledge share exactly the same functional properties of their corresponding representations.

The neurocognitive and philosophical accounts previously reviewed suggest a prominent link between the contents of declarative knowledge and the availability for verbalization. The way this link is encountered in the WPT literature is examined in the following section.

1.2.2. Verbal reports in the study of probabilistic category learning.

Despite the fact that, according to the previously reviewed accounts, items of declarative knowledge have been characterized by their ability to be verbally expressed, this feature had not been used in the experimental study of learning in the WPT, at least not in the early studies.

Knowlton et al. (1994) stated that declarative knowledge is "available to conscious recollection" (p. 106), but the method used in the study to assess explicit knowledge of the task was participants' answers on a multiple-choice questionnaire. Reber et al. (1996) stated that a property of the declarative memory system is "awareness that subjects have for what has been learned" (p. 861). Likewise, multiple-choice questions that did not require participants' verbal reports comprised their transfer task in order to assess flexibility of declarative knowledge. The same method was also used in the declarative memory task in Knowlton et al. (1996).

To the best of our knowledge, participants' verbal reports were first utilized in an experimental manner in the WPT literature in the strategy analysis introduced by Gluck et al. (2002). Participants were asked (see Section 1.1.4), after the categorization task, to verbally respond to an open-ended question and describe the strategy they had used to predict the weather during training. Their responses were used in order to formulate four strategies (see Fig. 1.5) that were thought to support performance in the task. Gluck et al. implied that a verbalizable strategy is mediated by the declarative memory system and they claimed, referring to the one-cue strategy, that a participant could "easily verbalize this strategy using declarative memory" (p.

416). An analogous assumption was made for nonverbalizable strategies being procedurally mediated: “weather prediction task has always been assumed to be learned in an implicit manner by use of procedural, nonverbalizable rules” (p. 414).

Yet, analysis of participants' patterns of performance according to strategy analysis was found to be uncorrelated to their verbal reports (as mentioned in section 1.1.4). This prevented Gluck et al. (2002) from drawing firm conclusions about brain structures mediating verbalizable versus nonverbalizable strategies. However, a potential counter-argument against the importance of the observed inconsistency between participants' verbal reports and the results of strategy analysis was offered by the same group of researchers. On a following elaboration on the method of strategy analysis, Meeter et al. (2006) reviewed the aforementioned inconsistency and commented that “participants reported retrospectively on their behaviour, and it is well known that participants during task performance may be aware of more than they will report after the task is completed (Ericsson & Simon, 1984)” (p. 237).

Although strategy analysis failed to provide clear evidence as to the memory systems underlying verbalizable versus nonverbalizable strategies, evidence on this issue had already been offered by the investigation on the activation of brain structures during the acquisition of knowledge in the WPT. Shohamy et al. (2004) linked the findings of both strategy analysis (Gluck et al., 2002) and neuroimaging (Poldrack et al., 2001) in the following way:

> One possible interpretation of these results is that the medial temporal lobes are involved early in learning (when participants might be investigating simple, easily verbalizable strategies), whereas basal ganglia involvement increases later in training (as subjects move to more optimal, less verbalizable strategies) (Shohamy et al., 2002, p. 684).

Thus, although the availability for verbal reports had not been used as an experimental measure in the early studies employing the WPT, this characteristic of declaratively acquired knowledge was introduced in the mathematical modelling of participants' behaviour. To the extent that the fitted strategies indeed reflect participants cognitive processes (see Gluck et al., 2002, for a discussion of this subject), it has been suggested (Shohamy et al., 2004, 2008) that early in learning performance is supported by declarative, simple, and verbalizable strategies, whereas in later stages of the task performance is supported by procedural, optimal, and nonverbalizable strategies.

1.3. Present Research

1.3.1. Purpose of present research

The purpose of the present research is to experimentally manipulate the contribution of the declarative memory system to the acquisition of knowledge in the WPT. Our own experience when participating in the task along with the reports of participants in the Gluck et al. (2002) study has led us to believe that WPT can be solved by some simple, verbalizable rules as opposed to the notion that the task is mediated by the procedural system.

Our effort was to render the task less declarative, or, according to accounts reviewed previously, less verbalizable, by using cues that participants would face difficulties in naming. Thus, we hypothesized that participants would be unable to develop declarative verbalizable rules, at least early in training, an effect which would be apparent in participants' early performance.

We hoped that an experimental manipulation of participants' ability to name the cues might provide a version of the WPT that could be faithful to the Knowlton et al. (1994) intention to “discourage the use of declarative memory” (p. 107), and thus become a more efficient way of assessing the incremental acquisition of cognitive skills. Also, we hoped that our research could contribute to the investigation of the assumed interaction between the declarative and the procedural memory systems, according to the MMS hypothesis.

1.3.2. Pilot studies and experimental manipulations

Our first effort was to implement a version of WPT using visual stimuli that would discourage the use of verbalizable strategies by using shades of the same colour. Inspired by the research on Universal Colour Names by Berlin & Kay (1969, as cited in Kay & Regier, 2003) and by the corresponding philosophical inquiries (Hardin, 2005), we constructed visual stimuli that we hoped would be discriminable to participants, but at the same time would be characterized by the same names. Our main idea was to use as cues cards coloured with shades of grey instead of cards containing geometric shapes. We hoped that participants would be able to perceptually discriminate each shade (and thus be able to improve in their categorization ability), but at the same time be unable to develop simple verbalizable rules such as “I predicted rain whenever I saw the triangle card” (Gluck et al., 2002, p. 411) since each card would be “grey” to them.

In particular, we used the findings of Sturges & Whitfield (1995) that administered a monolexemic naming task in order to define “consensus samples” (“a color sample (that) is described as *consensus* is one where all subjects named that sample consistently with the same color term”, p. 366) from the Munsell Color Space system. We used the study's consensus samples to create 4 shades of 4 colours (red, green, blue, and grey) and we implemented two tests: a two-alternative same-different task, in order to verify that the participants would be able to discriminate the shades of grey, and a free naming task, in order to verify that participants would use the same name when describing each shade of grey.

Our results revealed (data not shown) that although participants were almost 100% correct in discriminating the shades of grey, they would use names such as “light grey” or “dark grey” consistently from almost the first trials of the free naming task, in order to classify stimuli to different categories.

We concluded that humans (especially those of the female gender) have the ability to differentiate shades of colours easily by assigning different labels to them, and thus we focused on auditory stimuli inspired by the “absolute pitch” phenomenon.

It is well-studied that “most people are unable to name or place pitch values in consistent, well-defined categories” (Levitin & Rogers, 2005, p. 26). Some people, 1 in 10.000 (Ward, 1999, as cited in Levitin & Rogers) do posses the rare ability to label pitches without external reference but “most subjects do not possess pre-established labels for tones” (Galizio & Maron, 1976, p. 592).

It thus seemed to us that if we implemented an auditory version of the WPT using notes, or tones, as cues then the participants would face difficulties in naming the cues. Our pilot studies with colour shades indicated that in a categorization task, such as the WPT, it is extremely likely that participants would develop their own individual names for the cues, due to the repeated exposure to them, but we hoped that this would not happen with auditory tones as easily as with colour shades. Consequently, we were expecting that an auditory version of the WPT, with tones serving as cues, would discourage the use of simple verbalizable rules, at least in the first stages of the training, until participants had developed arbitrary names for the tones.

1.3.3. Assumptions, experimental design, and predictions.

The present research is based on several assumptions some of which have been previously mentioned, but in the following, for clarity, we will explicitly emphasize them and also describe this research's experimental design and prediction.

Our first assumption (see Section 1.2.2.) is that young healthy individuals approach the WPT early in training by simple, verbalizable rules mediated by the declarative memory system.

A second, quite central, assumption made in the present research is that verbalizability is an operating characteristic of declarative knowledge (Section 1.2.1.). Thus, we hypothesized, if we used in the WPT cues that participants would find difficult to name ("non-nameable" cues), then participants would be unable to develop, or rely on, declarative verbalizable rules in order to solve the task.

Our pilot studies suggested that participants might develop arbitrary names for the cues at some point in training. From that point on participants would be in position to develop verbalizable rules. Thus, it could be assumed that using non-nameable cues would prevent participants from engaging into verbalizable rules for a short, early period in training.

Thus, we concluded to a between-subjects experimental design. At first (Experiment 1) we administered the prototypical version of the WPT as typically implemented in the visual modality to a group of young healthy participants in order to make sure that all the details of the experimental set-up were correct and that we could replicate the findings of previous studies. Then (Experiment 2), we administered an auditory version of the task to a separate group of healthy individuals. In this condition, called the "nameable" condition, the cues used were animal sounds, which have been characterized as environmental, recognizable stimuli (Lewis, Wightman, Brefczynski, Phinney, Binder, & DeYoe, 2004), and we assumed that participants would find it easy to name them. Finally (Experiment 3), a third group of healthy individuals took part in another auditory version of the WPT in which we used computer-generated tones as cues. This condition was called the "non-nameable" condition since we thought that participants would face difficulties in naming the cues.

We hypothesized that participants in the non-nameable condition would be unable to develop, or rely on, simple verbalizable rules mediated by the declarative memory system early in training, at least till the point in training that they had developed arbitrary names for the cues. To the contrary, participants in the nameable condition would be able to develop and rely on verbalizable rules from the very beginning of the task.

Previous reports from studies with young healthy individuals in the typical visual version of the WPT indicate that their performance, early in the task, is close to chance levels. The learning curve of 30 participants (Experiment 2, Gluck et al., 2002), in blocks of 10 trials, is shown in Fig. 1.7. Meeter et al. (2006) commented on those results that "from a level close to chance, performance is increased to ~80% optimal answers" (p. 233).

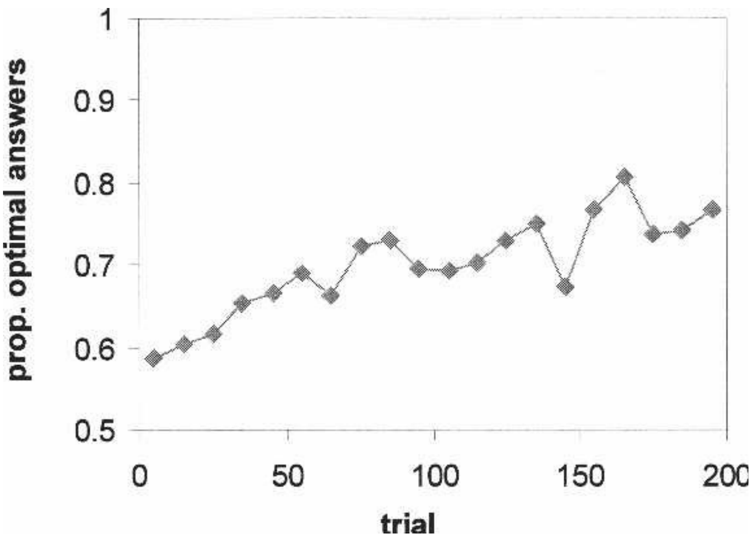

Figure 1.7. Performance of 30 participants in Experiment 2, Gluck et al., (2002). (Meeter et al., 2006, p. 234).

If the same pattern of performance is also exhibited in an auditory version of the WPT, then, it could be assumed that performance of participants in the nameable condition (supported by verbalizable rules early in training) could be at chance levels at the beginning of the task. According to our assumption though, early performance of participants in the non-nameable condition could not to be supported by verbalizable rules till the point in training that those participants had developed arbitrary names for the cues. So, our prediction is that performance in the non-nameable condition would be at chance levels for a longer period compared to the nameable condition.

A possible argument against the aforementioned prediction is that performance of participants in the non-nameable condition could be mediated by the procedural memory system, since the declarative memory system is discouraged and there has been suggested a competitive interaction between the two memory systems (Poldrack et al, 2001). Yet, the cortico-hippocampal model of category learning (Gluck & Myers, 1993) suggested an involvement of basal ganglia later in the task. Moreover, procedural knowledge, able to support high levels of performance, is thought to be incrementally acquired over the course of many trials (Shohamy et al., 2008). So, even if the procedural memory system is activated early in training, it is improbable to support better-than chance performance early in the task. Thus, a purported early engagement of the procedural memory system is assumed not to mask the predicted longer near-chance period in performance of participants in the non-nameable condition, relative to the nameable condition.

The experiments conducted and the corresponding results are discussed in the following sections.

Part 2
Experiments

2.1. Experiment 1.

2.1.1. Method

2.1.1.1. Participants and apparatus

Twenty students (4 male) of the Psychology Department, Panteion University, Greece, took part in Experiment 1, rewarded with course credit for their participation. Their mean age was 20.4 yr (SD = 4.65) and all of them (as well as participants in Experiments 2 and 3) reported normal hearing and normal or corrected-to-normal vision and no history of neurological illness.

The experiment was programmed in DMDX Display Software (Forster & Forster, 2003) and it was run on a Toshiba (Satellite A500) laptop computer. Participants wore a set of Panasonic (DJ 100) headphones throughout the experiment.

2.1.1.2. Materials

In Experiment 1 we replicated the WPT as described by Knowlton et al. (1994), using the "newer" structure of the experiment, i.e., the modified probabilities of Gluck et al. (2002).

The stimuli were a set of four cards ("cues"), each containing a different geometric shape (squares, diamonds, circles, and triangles, see Fig. 1.4). Participants completed a set of 200 trials, on each of which they were presented with one of 14 possible cue combinations ("patterns"). The frequency of each pattern along with the frequencies of the two outcomes ("sun" and "rain") is shown in Table 1, in which it can be seen that the two outcomes occurred equally often throughout the experiment (Gluck et al., 2002). Each cue is independently associated with an outcome with a fixed probability. This probability, P(outcome|cue present), can be calculated from Table 2 (as described by Shohamy et al., 2004). For example, cue 1 is present in patterns H to N, which in total appear in 100 trials of the experiment. In these 100 trials the outcome of sun occurs 20 times and the outcome of rain occurs 80 times Thus, cue 1 is associated with sun with probability 20 ÷ 100 = 0.2 and with rain with probability 0.8. Likewise, it can be calculated that cues 2, 3, and 4 predict sun with probabilities 0.4, 0.6, and 0.8 respectively. The assignment of shape type (squares, diamonds, etc.) to associative strength (C1, C2, etc.) was counterbalanced across participants, while the relative position of each card on the screen was held constant for a given pattern and a given participant.

Table 1

Pattern and Outcome Frequencies Used in Experiments 1 and 2 as defined in Gluck et al., 2002.

	Cue						
Pattern	1	2	3	4	Sun	Rain	Total
A	0	0	0	1	17	2	19
B	0	0	1	0	7	2	9
C	0	0	1	1	24	2	26
D	0	1	0	0	2	7	9
E	0	1	0	1	10	2	12
F	0	1	1	0	3	3	6
G	0	1	1	1	17	2	19
H	1	0	0	0	2	17	19
I	1	0	0	1	3	3	6
J	1	0	1	0	2	10	12
K	1	0	1	1	5	4	9
L	1	1	0	0	2	24	26
M	1	1	0	1	4	5	9
N	1	1	1	0	2	17	19
Total					100	100	200

The order of trials was random but fixed for all participants, as described by most studies employing the WPT (e.g., Gluck et al., 2002; Knowlton et al., 1994; but see Lagnado et al., 2006, Newell et al., 2007; Price, 2009; Wilkinson et al., 2008). The order of trials and the corresponding outcome on each one of them were kindly provided to us by Prof. Martijn Meeter (personal communication, December 4, 2009) and can be seen in the Appendix (along with our own devised order for trials 201 to 302 implemented in Experiment 3).

2.1.1.3. Procedure

Experiments were conducted in quiet surroundings at the University library. Participants were seated at a comfortable viewing distance from the laptop screen, and wore a set of headphones.

Participants were informed that they would take part in a learning experiment, but they were not informed of the probabilistic nature of the task. All participants executed five trials of the WPT before the actual experiment, in order to become familiar with the experimental environment and in order to adjust the sound volume to their own comfortable levels. In this "familiarization" phase all the instructions were almost identical to the ones given in the actual categorization task (as described below), but the cues were cards with different colours (red, green, blue, yellow) to prevent any kind of learning of the cue-outcome contingencies. With very few exceptions (in which the familiarization phase was repeated once more) all participants by the end of this phase were able to respond by pressing the appropriate keys on the keyboard within the required time and they then took part in the actual task.

At the start of the categorization task, participants were given on the computer screen the following instructions (which were an adaptation of those described by Lagnado et al., 2006):

> In this experiment you will be playing a game in which you will be the weather forecaster. In each trial you will be seeing one to three cards. You have to decide whether the combination of cards present predicts "sun" or "rain", and press on the keyboard the right shift for "sun" or the left shift for "rain". At first, you will have to guess, but you will gradually become better in deciding. A yellow bar on the screen will rise in height with every correct answer, and will fall with every wrong answer. In total you will see 200 combinations of cards, and every 50 combinations you will be able to take a short break. Whenever you feel ready to start, please press the "space" key.

At the start of each trial a pattern appeared at the center of the screen, and above it appeared one icon of the sun on the right side and one icon of a raining cloud on the left side. The computer recorded the participant's response, and the actual outcome (one of the two aforementioned icons) appeared above the pattern for 2 s along with the feedback. The feedback consisted of a frowning smiley and a low tone (frequency: 500 Hz, duration: 0.1 s) when the participant predicted the wrong outcome, or a happy smiley and a high tone (frequency: 1000 Hz, duration: 0.1 s) in the opposite case. If the participant did not respond within 2 s, a “Please respond now” prompt appeared, and if he/she did not respond within the next 3 s either, the trial was terminated and the participant was scored as having made the wrong response on that trial. Following Knowlton et al. (1994), there was a yellow bar on a black background present on the right side of the screen. At the start of the experiment the bar had a height of 200 pixels and it could rise or fall by one pixel on each trial according to the feedback provided to the participant on his/her response. This change in the yellow bar was apparent on the beginning of the next trial, in order not to capture the participant’s attention but provide a rough (but not accurate, see Data analysis below) estimate of his/her performance. The height of the bar would remain unchanged in case the participant made no response on a trial, and at the end of the experiment its height could range from 0 (all responds receiving negative feedback) to 400 pixels (all responds receiving positive feedback). The intertrial interval was half a second, after which the next trial begun. As mentioned in the instructions, three short breaks (whose duration was determined by the participants by the pressing of the space key) intervened every 50 trials.

The duration of the experiment was on average 15 to 20 minutes.

2.1.1.4. Data analysis

Participants' performance was measured in terms of “optimal responses” (Knowlton et al., 1994). A participant's response was thought to be optimal and it was marked as “correct” if he/she chose the outcome more often associated with that pattern, whereas it was marked as “wrong” in the opposite case. Thus, a participant could have been marked as having made the correct response on a given trial, although the feedback he/she received could indicate a wrong decision. As mentioned previously, a lack of response was marked as wrong response. Moreover, two of the patterns presented in the experiment, patterns F and I (appearing in overall 12 trials, as can be seen in Table 1) predict “sun” or “rain” with equal probabilities. Thus, no optimal response could be defined for these 12 trials and participants' responses on them were not included in the analyses. Experimental data were preprocessed with the azk2txt software (distributed along with CheckVocal software, Protopapas, 2007) and were analysed using SPSS software; additional analyses were conducted with R (Baayen, 2008).

2.1.2. Results

Over all 200 trials, participants averaged 76.81% optimal responses (SD = 8.98%). Fig. 2.1 shows the percentage optimal responding in blocks of 50 trials. We conducted a repeated-measures ANOVA, and there was a significant within-subjects effect of block on performance: $F(3, 57) = 3.821$, $p = 0.015$. Moreover, polynomial contrasts analysis revealed a significant linear trend: $F(1, 19) = 10.1$, $p = 0.005$. These results indicate that participants' performance in the categorization task improved across blocks.

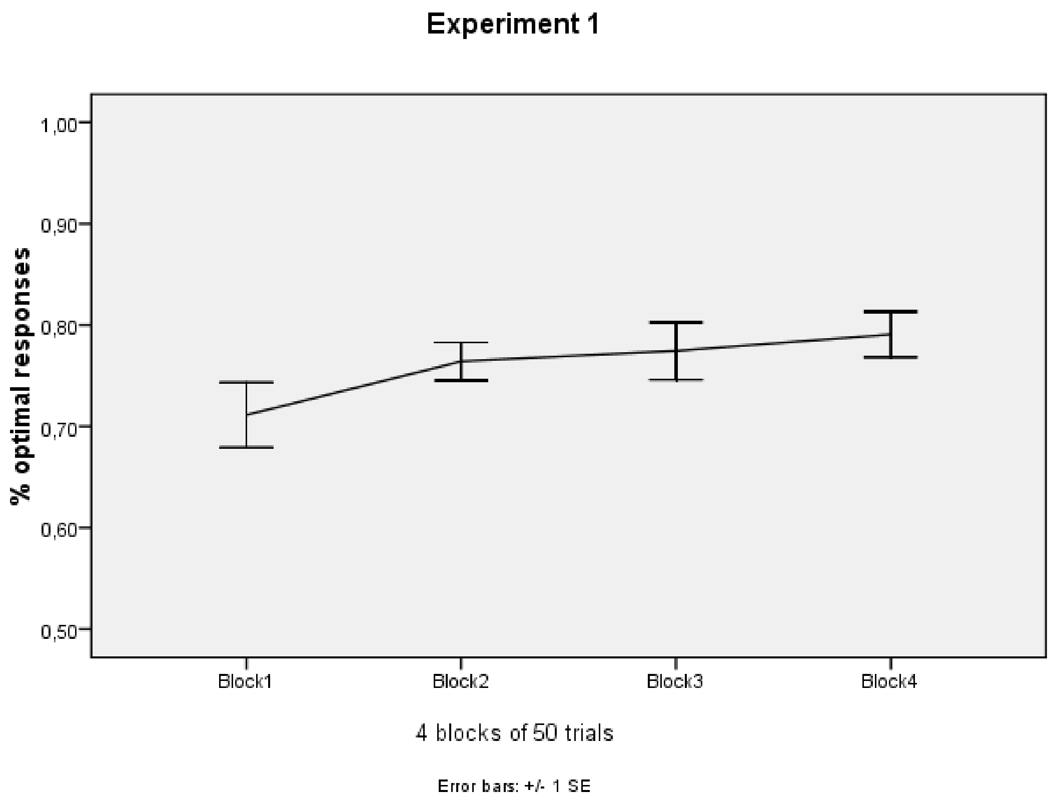

Figure 2.1. Experiment 1 data: learning curve over all 200 trials analysed in 4 blocks of 50 trials.

2.1.3. Discussion

Our results are in line with the performance levels and the shape of the learning curve observed in previous studies of the WPT that recruited young healthy individuals (Gluck et al., 2002). Although most of our participants in Experiment 1 (and also in the following experiments) were female, we believe this did not pose any problems for the purposes of our research, based on previous findings that revealed no effect of gender on performance (Gluck et al.). The main reason we implemented this replication of WPT was to ensure that the experimental setup and all the details of the experiment (e.g., the order of trials and the analysis of data) that we utilized were correct. Our results suggested that a potential lack of evidence of learning in the auditory modality should not be attributed to the experimental set-up.

2.2. Experiment 2

2.2.1. Methods

Two groups of healthy individuals participated in Experiment 2. The first group (mean age: 19.65, SD = 2.48) comprised 20 psychology students (4 male) of Panteion University. The second group (3 male, mean age 23.55, SD = 6.4) comprised 15 students of the same department (receiving course credit for their participation) and 5 students of the Basic and Applied Cognitive Science master's degree program of the National University of Athens; none of them had any previous experience with the WPT.

The probabilistic structure of the experiment, the apparatus, and the analysis of data in the categorization task were identical to those in Experiment 1. Differences in the stimuli, the experimental procedure (a post-training test task was administered to participants in Experiments 2 and 3) and the analysis of data in the test task are described below.

2.2.1.1. Materials

The four cues for Experiment 2 were selected from a set of 10 animal sounds that had been used recently in a task requiring their explicit identification (Seitz et al., 2010). The duration of the sounds was 500 ms and they were equated, to the extent possible, in intensity. We conducted a two-alternative forced-choice free association pilot experiment (data not shown), in which 8 university colleagues listened to all 10 sounds and had to respond by selecting either “sun” or “rain”. We excluded all the sounds that were unanimously associated with either outcome, and chose the sounds that exhibited maximum variability in their association with the notion of good or bad weather, in order to minimize the contribution of any accidental learning cues to performance (e.g. the bird sound always yielded the “sun” response, and was therefore excluded). Using this procedure the sounds of four animals that were selected as cues for Experiment 2 were: the sound of an elephant, a dog, a cricket, and a cow; their spectrograms can be seen in Fig. 2.2.

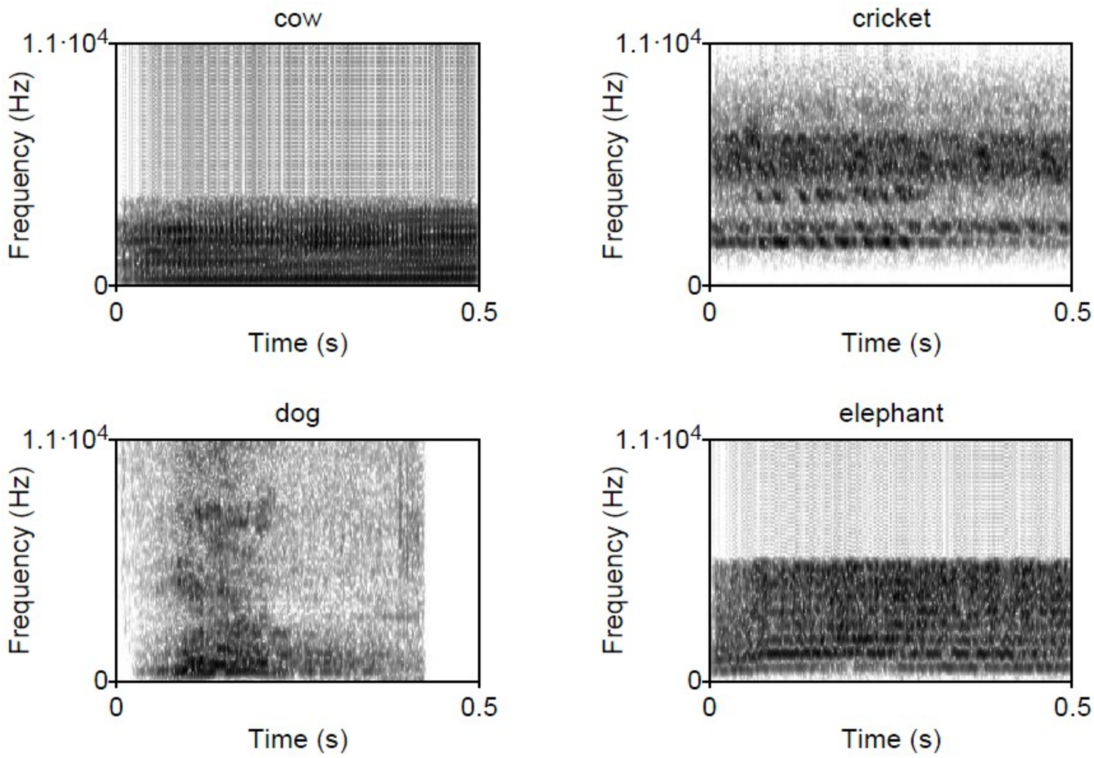

Figure 2.2. Spectrograms of the four sound by Seitz et al. (2010) that served as cues in Experiment 2.

2.2.1.2. Procedure

Instructions to participants in Experiment 2 were almost identical to those used in Experiment 1, adjusted where needed to reflect the auditory nature of the cues.

The sequence of events in a trial in this auditory version of the WPT was not the same as in the prototypical WPT. We chose to present the auditory cues sequentially, assuming that participants process the complex visual stimuli in the original version also sequentially, based on reviews of visual attention mechanisms (e.g., Egeth & Yantis, 1997). Hence, the duration of “patterns” of auditory cues was not the same for all patterns, but depended on the number of cues comprising a pattern, and ranged from 2 s (1-cue pattern) to 5 s (3-cue pattern). The sequence of events in a 2-cue pattern is shown in Fig. 2.3, and it can be seen that there was an interval of 1 sec between cues. After having heard the auditory pattern (the end of which was signalled by the presentation of the sun and rain icons), participants were given 5 s to provide their response, as in the original version. The same prompt as in Experiment 1 was used if participants had not provided a response within 2s, and after their response the actual outcome remained on the screen for 2 sec, along with the auditory and visual feedback (positive or negative) that was also used in Experiment 1.

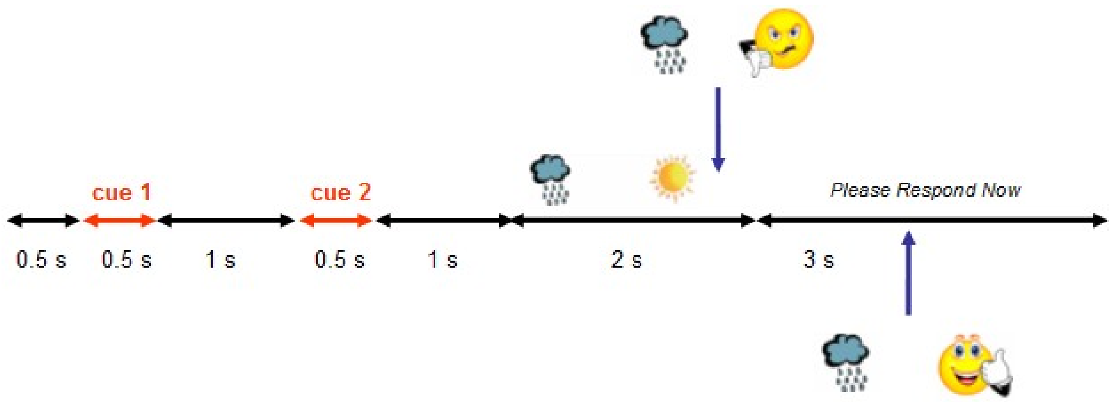

Figure 2.3. Timing of events in a 2-cue pattern trial in Experiment 2. Blue arrows designate two potential alternative responses by participants.

This difference in the sequence of events in a trial introduced an important discrepancy between the original version of the WPT and this auditory version, which is thought to have implications on the nature and difficulty of the task. Participants in the prototypical version of the WPT were presented with the cues and the actual outcome during feedback on each trial, but in our auditory version they were presented with only the actual outcome. The former, prototypical form of feedback is thought to increase the contribution of the declarative memory system in learning (e.g., as assumed for the paired-associate task in Experiment 1 of the Poldrack et al., 2001 study). Consequently, the version of the WPT implemented in Experiment 2 (and 3) is thought to not encourage the engagement of the declarative memory system compared to the prototypical version. The difference in the modality between Experiments 1 and 2 of the present research does not allow for an isolation of the effect of the form of feedback used. This discrepancy in the from of feedback is further discussed in Part 3, along with a prediction and a suggested experimental manipulation.

We also speculated that in order to solve the task, participants in our auditory version, would have to keep in their working memory (or echoic memory) the cues previously heard in order to learn the cue-outcome associations. So, presumably, the difficulty of this auditory version of the task would be increased compared to the original WPT, since it requires additional cognitive resources.

As in Experiment 1, there was also a “familiarization” phase at the start of the Experiment 2, and the sounds used in these five trials were the complex tones that served as cues in Experiment 3.

Finally, participants in Experiment 2 (and 3), after having completed the categorization task, also took part in a test task. They were given a questionnaire and were asked to write down their answers to the following two questions:

Question 1:
If the weather turns out to be fine, which sound is it most likely that you've heard before?

Question 2:
If the weather turns out to be rainy, which sound is it most likely that you've heard before?

We consider this questionnaire to be a verbal variant of the Cue-Selection task used in previous studies in order to provide an estimate of participants' declarative knowledge of cue outcome-associations (Foerde et al., 2006, 2007; Reber et al., 1996). We utilized this test not with the same purpose, but in order to assess the verbal behaviour of the participants when asked to provide names for the cues of the categorization task.

The duration of the Experiment 2 (familiarization phase, categorization task, and test task) was on average 30 to 35 minutes.

2.2.1.3. Data analysis

Data from the test task were quantified and two measures were extracted.

First, we measured participants' performance on cue-selection, i.e., we checked if they were aware of the cues that were highly predictive of sun (Question 1) and rain (Question 2). Following Foerde et al. (2007), the answer on the first question was scored with 4 if participants provided the name of the sound that was highly predictive of sun, with 3 if the sound was less predictive of sun, with 2 if it was less predictive of rain, with 1 if it was highly predictive of rain, and with 0 if no answer was given. The analogous quantification was applied to the answer in the second question, and the two scores were added; the sum comprised each participant's Cue-Selection performance that could range from 0 to 8 with chance performance being 4.087.

Second, we extracted participants' performance on successful naming. Each answer was scored with 4 if the name given for an animal sound was correct (i.e., if they named the cow sound as “cow”) and with 0 if the name was wrong (e.g., if they named the cow sound as “sheep”), regardless of the sound's associative strength (i.e., regardless of the score in the previous measure). The two scores were also added and the sum comprised the participant's performance in Cue-Naming that could range from 0 to 8. with chance performance being 4.

2.2.2. Results

2.2.2.1. Categorization Task

As mentioned previously, two groups of 20 participants each were administered the categorization task.

The first group's performance over all 200 trials averaged 73.38% (SD = 11.37) and the learning curve is depicted in Fig. 2.4. Performance increased gradually across the 3 first blocks of 50 trials from an average of 70.82% optimal responses in the first block to 75.00% optimal responses in the third block. There was no further improvement in performance since in the fourth block performance was on average exactly the same as in the third block: 75%. A t-test revealed that performance on the last block was significantly above chance $t(19) = 8.841$, $p < 0.001$, but a repeated-measures ANOVA revealed no significant effect of block on performance ($F(3, 57) = 1.29$, $p = 0.287$), which is usually taken to be indicative of gradual learning in the WPT (Gluck et al., 2002; Knowlton et al., 1994; but see Hopkins et al., 2004, reporting no effect of trial block in performance, and almost U-shaped learning curves of both control and hypoxic participants). Moreover, polynomial contrasts analysis revealed that the linear trend was not significant, $F(1, 19) = 2.265$, $p = 0.149$. The results of our analyses could not firmly indicate gradual learning, but by the time of the analysis in question the data from Experiment 3 indicating gradual learning were available to us. Thus, we decided to administer Experiment 2 to a second group of 20 participants, to test whether the lack of gradual learning was replicable or due to chance.

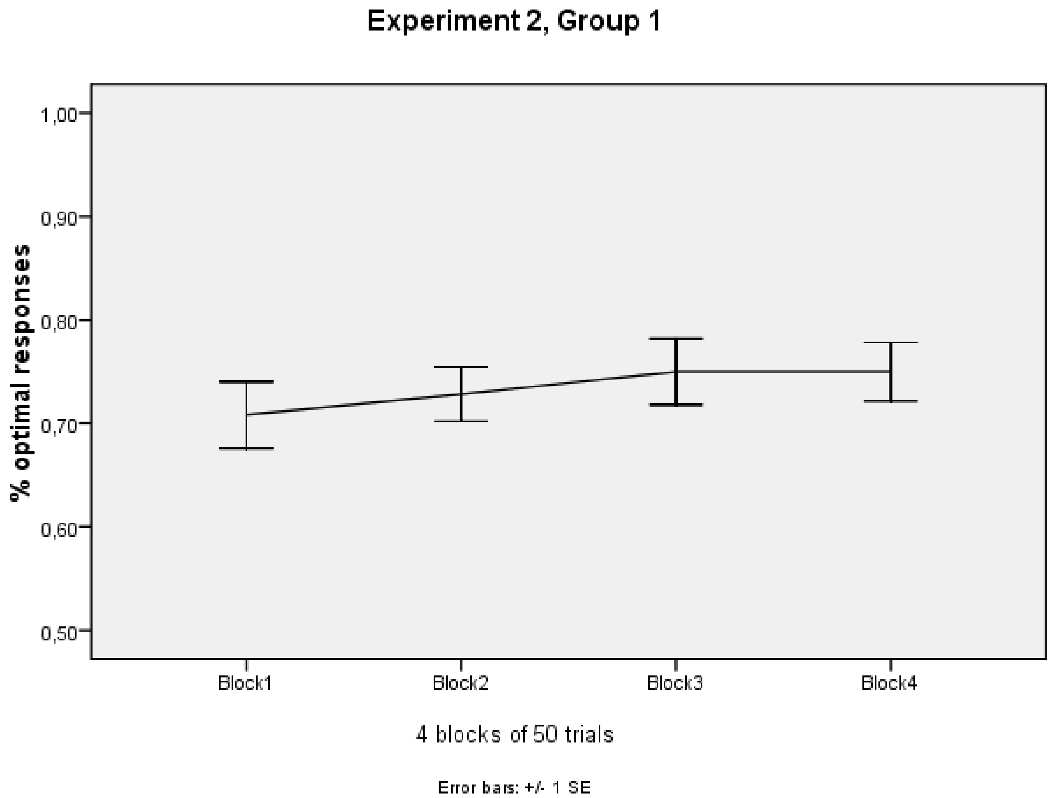

Figure 2.4. Performance of the 1st group of participants in the categorization task, Experiment 2.

The second group over all 200 trials averaged 75.24% optimal responses (SD = 6.4%) and the learning curve is shown in Fig. 2.5. This time participants' performance increased gradually across the four blocks (from an average of 70.82% optimal responses in the first block to 81.67% in the last block) which was confirmed by a repeated-measures ANOVA revealing a significant effect of block on performance, $F(3, 57) = 5.137$, $p=0.003$ (Mauchly's test revealed that the assumption for sphericity had been violated, but Greenhouse-Geisser corrections on degrees of freedom also revealed a significant effect; for the sake of simplicity these corrections are not reported). Polynomial contrasts analysis also revealed a significant linear trend, $F(1, 19) = 7.424$, $p = 0.013$.

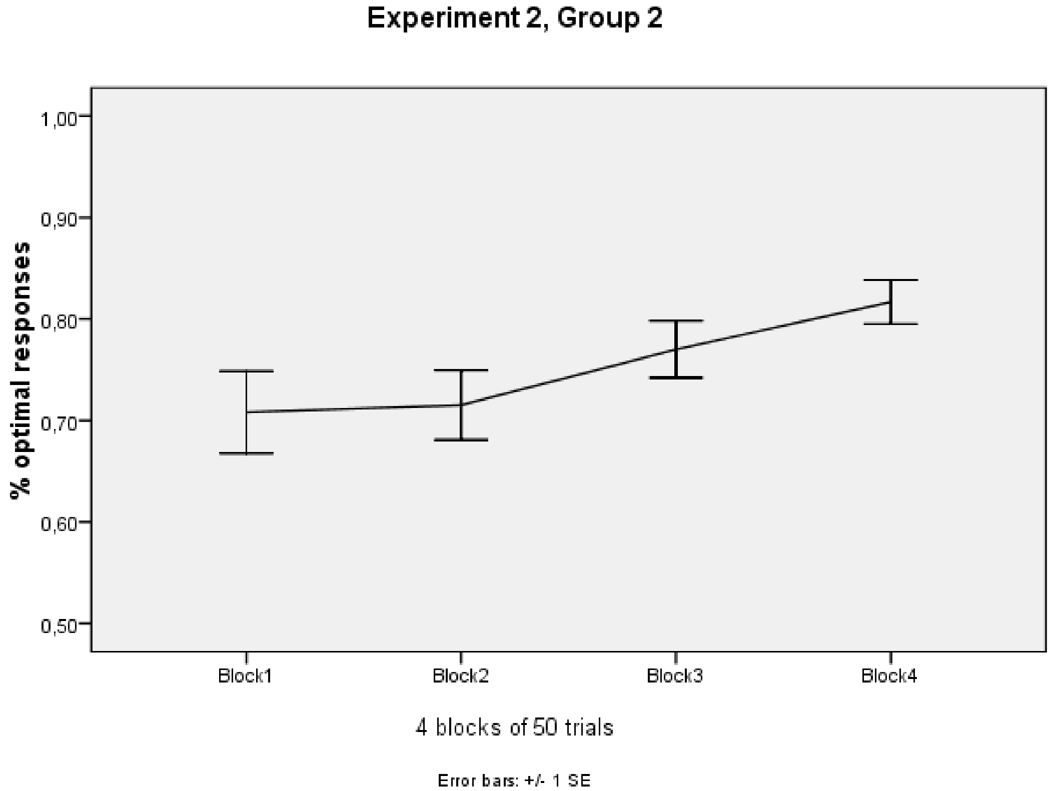

Figure 2.5. Performance of the 2nd group of participants in the categorization task, Experiment 2.

This discrepancy in our results was puzzling and intriguing so we conducted some more analyses. By excluding data from two individuals from the 1st group who seemed to have given up in their effort to solve the categorization task (as indicated by their individual learning curves that showed decreases in performance throughout the four blocks) the data from 18 participants showed the desired effect, namely a significant effect of block on performance. Nevertheless, Prof. Martijn Meeter (personal communication, May 25, 2010) suggested that it is not legitimate to exclude participants that seem to have given up on their mediocre performance, so our next effort was to pool and analyse data from the two groups in order to increase the statistical power of our analysis. This analysis of behavioural data from 40 participants revealed a significant effect of block on performance, $F(3, 117) = 5.764$, $p = 0.001$ (Greenhouse-Geisser corrections are not reported, as previously), and a significant linear trend, $F(1, 39) = 9.37$, $p = 0.004$, that resemble the characteristics of the second group's performance. Also, it is worth mentioning that the observed discrepancy in performance between the two groups is due to differences towards the end of the experiment. We analysed the data of both groups in 20 blocks of 10 trials, and performance in the first 5 blocks, which includes and exceeds the period of interest for our prediction (as described in section 1.3.3.), is highly similar in both groups as suggested by Fig. 2.6 and also by a two-way ANOVA (with group taken to be the between-subjects factor and trial block the within-subjects factor) of data in these first 5 blocks of 10 trials. The analysis revealed no interaction of group by block, $F(4, 152) = 0.519$, $p = 0.722$, and no main effect of group on performance, $F(1, 38) = 0.001$, $p = 0.973$. Thus, we concluded that lack of evidence of gradual learning in the first group was due to chance and that, overall, participants seem to gradually improve in performance across trials in Experiment 2. In the rest of this report we will be depicting the data of the second group of 20 participants as representative of healthy individuals' behaviour in Experiment 2 (clearly stated in the opposite case), since these data bear the characteristics of the pooled data of 40 participants, and since data from both groups lead to the same results when examining the main prediction of the present research (as described in section 2.4.1.).

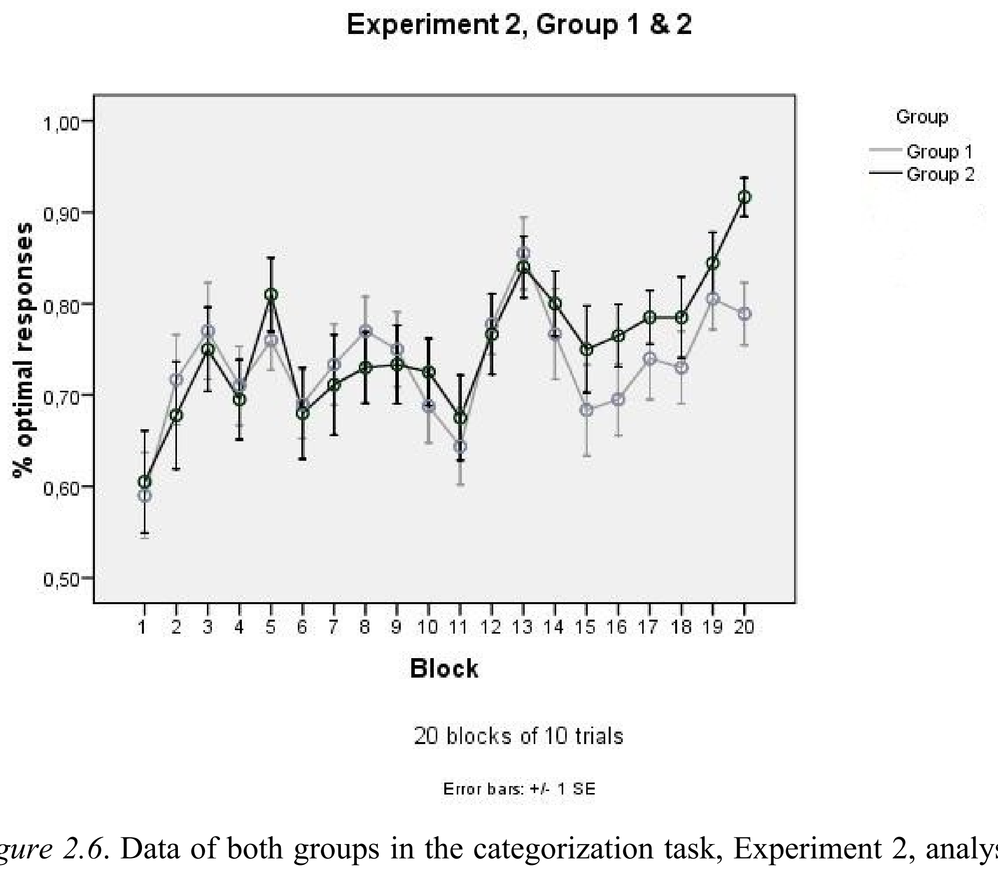

Figure 2.6. Data of both groups in the categorization task, Experiment 2, analysed in blocks of 10 trials.

2.2.2.2. Test Task

Participants in the test task of Experiment 2 (group 2) were willing to fill in the questionnaire administered to them.

We applied an arbitrary success criterion of 65% optimal responses in the final block of 50 trials in the categorization task (introduced by Gluck et al., 2002 as a criterion of assumed learning applied, though, to each participants' performance over all 200 trials), to ensure, to the extent possible, that replies were not random and that they reflected participants' knowledge acquired throughout training. Analysis of the replies (as described in section 2.2.1.3) of 19 participants that met the arbitrary criterion revealed that they were aware of the cue-outcome contingencies, since their score in Cue-Selection was 7.37 (SD = 0.831) with the maximum score being 8. These results are consistent with results of previous studies indicating a high score in cue-selection (assessed, though, by means of different questions in nature and quantity) when the WPT is administered without a distracting secondary task (Foerde et al., 2006, 2007), and with studies reporting that participants, when probed, "on average, tend to provide relative accurate estimates of cue-outcome associations" (Gluck et al., 2002, p. 412).

Participants were also successful in providing the correct names for the animal sounds that served as cues in Experiment 2, since only 4 of them (out of 19) provided wrong names for one of the two sounds probed for. Three more participants named the cricket sound as "sound of the night", but we scored that response as correct, so the score in Cue-Naming was 7.16 (SD = 1.675) with maximum score being 8. Participants' scores in both measures are depicted in Fig. 2.7.

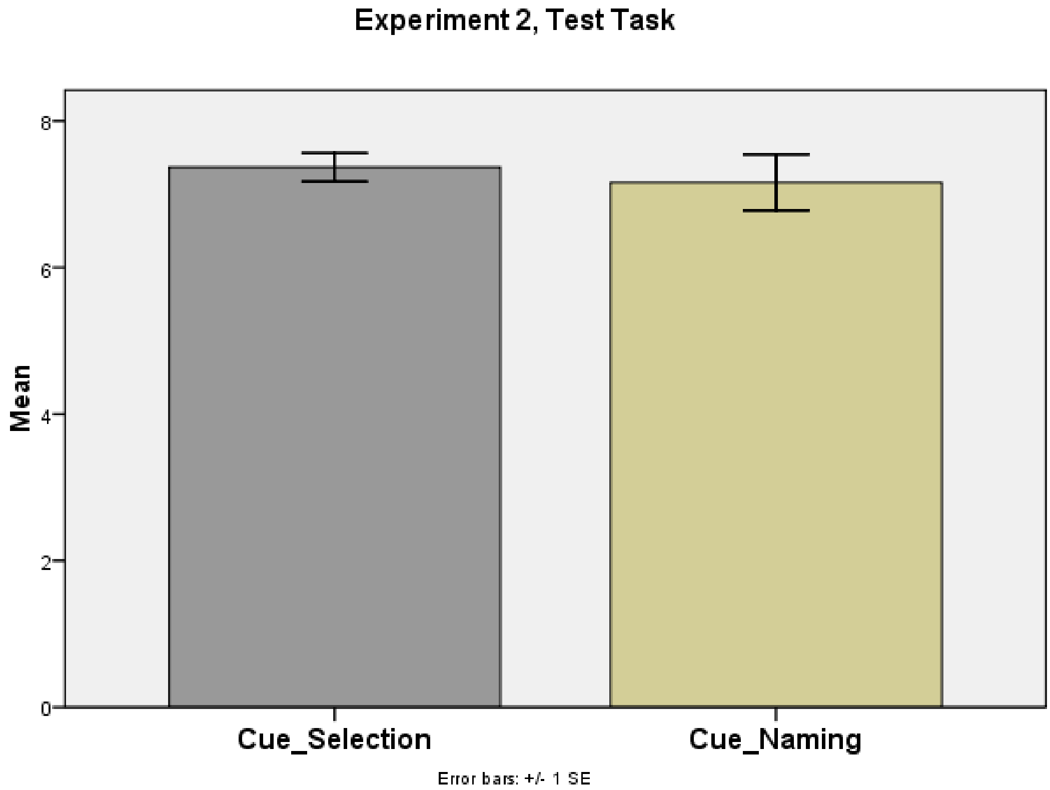

Figure 2.7. Performance of participants in the test task, Experiment 2

2.2.3. Discussion

Results from the categorization task in Experiment 2 indicate that there is evidence of gradual learning in this auditory version of the WPT, and that despite of the assumed increased difficulty of the task (as discussed in section 2.2.1.2.) participants improved in their ability to predict the outcome across training trials. To the best of our knowledge, this result, with respect to the auditory modality, has not been reported previously.

Moreover, results from the post-training test task (particularly participants' cue-naming score) indicate that the animal sounds that served as cues in Experiment 2 were indeed recognizable and easy to name, so the characterization of this condition as “nameable” condition has been experimentally justified.

2.3. Experiment 3

2.3.1. Method

Twenty psychology students (2 male) of Panteion University (mean age: 20.25, SD = 3.46) participated in Experiment 3 and received course credit. Apparatus and data analysis in the categorization task were identical to those in Experiment 2. Differences in the stimuli, procedure, probabilistic structure, and data analysis of the test task are described below.

2.3.1.1. Materials

In Experiment 3 we tried to implement the “non-nameable” condition (as described in section 1.3.3), thus we used computer-generated complex tones as cues. These tones were similar to tones used by Holt & Lotto (2006) to assess participants' cue-weighting in an auditory categorization task. 192 frequency-modulated tones were generated at Carnegie Mellon University and they were made available to us by Prof. Lori Holt's student, Sung-Joo Lim (personal communication, January 8, 2010), defining 4 categories in the center frequency by modulation frequency two-dimensional auditory space. The position of all the tones in this space is shown in Fig. 2.8, along with the 4 tones (depicted in red) selected to serve as cues in Experiment 3. As can be seen in Fig. 2.8, we selected the tones so that their relative distance would be to the extent possible equal, in order to equate the potential perceptual discriminability between each pair of cues. The characteristics of the four tones are shown in Table 2 while their duration was 300 ms.

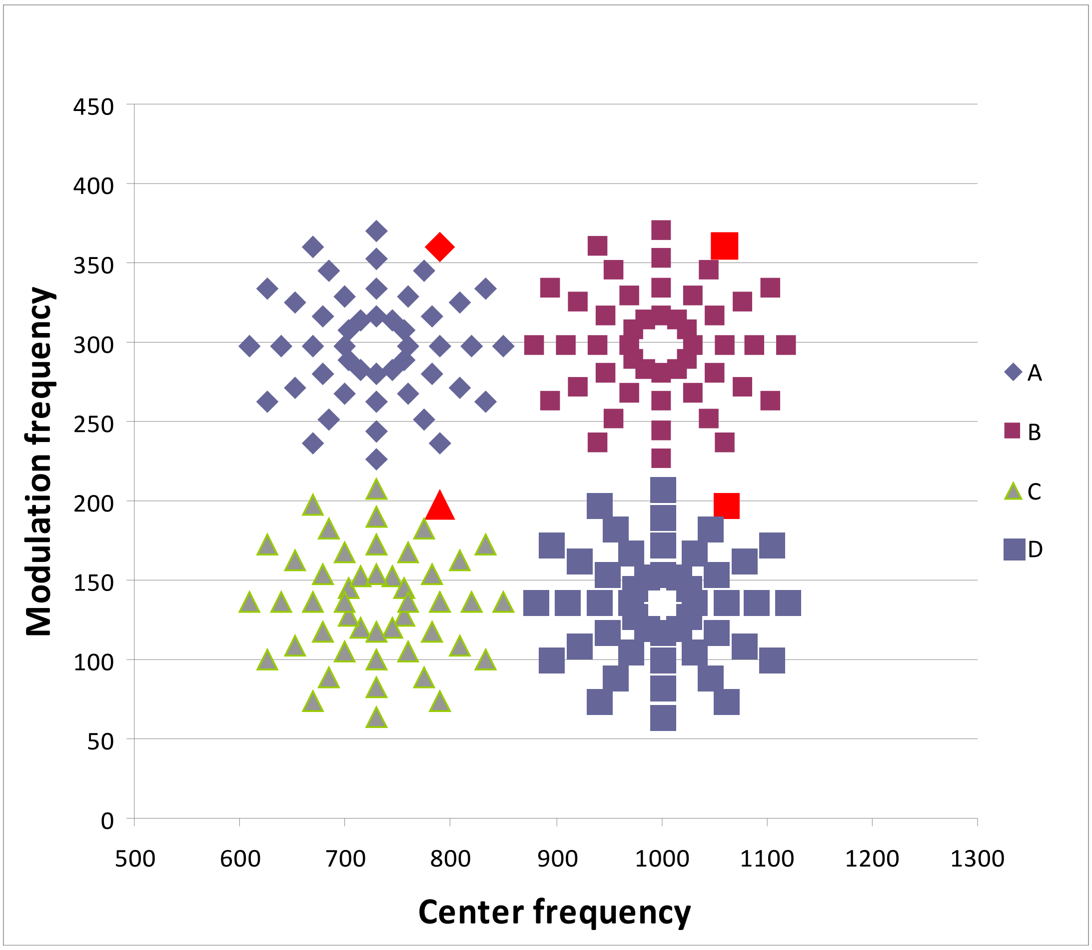

Figure 2.8. Distribution of 192 complex tones in the center frequency by modulation frequency acoustic space (Sung-Joo Lim, personal communication), and the four categories in this space, each comprised by 48 tones. Tones depicted in red are the four tones selected to serve as cues in experiment 3.

Table 2

Center and Modulation Frequency of the Four Frequency-Modulated Tones that Served as Cues in Experiment 3.

	Center frequency (Hz)	Modulation frequency (Hz)
Tone 1	790	360
Tone 2	1060	360
Tone 3	790	198
Tone 4	1060	198

For Experiment 3 we predicted a longer (compared to Experiment 2) near-chance-performance period early in the task as a result of the non-nameability of the cues used. Assuming that participants in Experiment 3 might develop arbitrary names for the complex tones at some point in the task, it is likely that—after that point—the learning rate observed would be the same as in Experiment 2 (since both the procedural and the declarative memory system may contribute to performance). Therefore, participants in Experiment 3 might require a protracted learning period in order to achieve the same levels of performance by the end of the task as participants in Experiment 2. In order to test this possibility we expanded the training of participants in Experiment 3 beyond 200 trials

We used the probabilistic structure of a 102-trial version of the WPT implemented by Newell et al. (2007) along with the 200-trial structure of Gluck et al. (2002) and the unification led to a 302-trial version whose probabilistic structure is shown in Table 3. Attention was given so that the cue-outcome associations resulting from the relative frequencies of the structure of the experiment would remain unaltered. Indeed, as can be calculated from Table 3 (method discussed previously) the cue-outcome associations are characterized by the same probabilities as in Experiments 1 and 2, namely cues 1, 2, 3 and 4 predict sun with probability 0.2, 0.4, 06, and 0.8 respectively. Also, the two outcomes occur equally often throughout the experiment, 151 times each, and thus chance performance remains at 50%.

Table 3

Pattern and Outcome Frequencies Used in Experiment 3

	Cue						
Pattern	1	2	3	4	Sun	Rain	Total
A	0	0	0	1	26	3	29
B	0	0	1	0	11	3	14
C	0	0	1	1	36	3	39
D	0	1	0	0	3	11	14
E	0	1	0	1	15	3	18
F	0	1	1	0	5	5	10
G	0	1	1	1	25	3	28
H	1	0	0	0	3	26	29
I	1	0	0	1	5	5	10
J	1	0	1	0	3	15	18
K	1	0	1	1	7	6	13
L	1	1	0	0	3	36	39
M	1	1	0	1	6	7	13
N	1	1	1	0	3	25	28
Total					151	151	302

The order of trials used in Experiment 3 was identical to the one used in Experiments 1 and 2 for the first 200 trials. For trials 201 to 302 we devised a random but fixed order of trials and of their corresponding actual outcomes for all participants. The order can be seen in the Appendix, and provision was taken so that the same outcome is not presented more than three times in consecutive trials. Also, following Knowlton et al. (1994) the same cue pattern is never presented in two successive trials. Overall we tried to distribute uniformly throughout the task trials incorporating the same pattern and moreover the same effort was made with respect to two specific kinds of trial: (1) trials that yielded a "surprising", i.e., the less probable, outcome (e.g., pattern A predicts sun most of the times, but in 3 out of 29 trials the actual outcome is rain, a fact which caused frustration to participants in Experiment 1 and thus we consider it to affect learning, see also Hopkins et al., 2004, for a note on participants' frustration, and Section 3.3. for further discussion), and (2) trials that are excluded from the analysis of behavioural data, i.e. trials in which patterns F and I are presented (for which no optimal response can be defined). In Fig. 2.9 the distribution of these two kinds of trials throughout the entire experiment is shown, which allows for a comparison between the standard distribution derived from the order of trials used in previous studies for the first 200 trials (e.g., Gluck et al., 2002) and our own distribution for the last 102 trials.

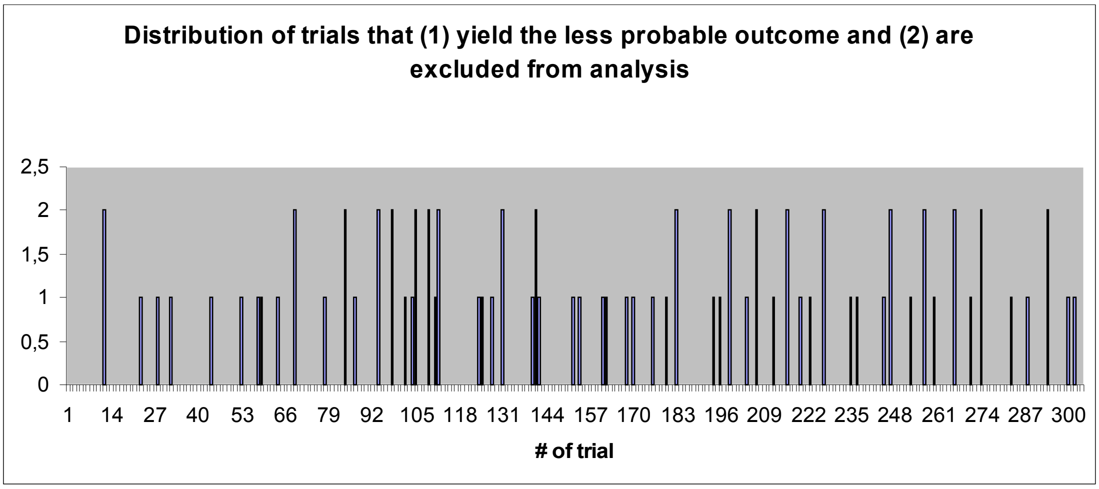

Figure 2.9. Trials that yield the less probable outcome (bars with height set to 1) and trials that are excluded from the analysis (bars with height set to 2) throughout the 302-trial version of the WPT implemented in Experiment 3.

2.3.1.2. Procedure

Instructions to participants in Experiment 3 were identical to the ones used in Experiment 2, with the sole difference being the reference to 302 instead of 200 combinations of sounds presented.

At the start of the experimental procedure there was a familiarization phase administered to participants and the sounds in these five trials were the animal sounds that served as cues in Experiment 2.

Also, there were two more short breaks (resulting to a total of 5 breaks) added to the categorization task.

The duration of Experiment 3 (familiarization phase, categorization task, test task) was on average 40 to 45 minutes.

2.3.1.3. Data Analysis

Data from the test task could not be quantified by the same method as in Experiment 2 (for reasons explained in Section 2.3.2.2.). Instead, the participants' replies to the questionnaires were collected and shown.

2.3.2. Results

2.3.2.1 Categorization Task

Over all 302 trials participants averaged 66.28% optimal responses (SD 8.47%). Fig. 2.10 shows the percent optimal responses in blocks of 50 trials. Participants performance improved from 54.49% in the first block to 73.19% in the final block, and a repeated measures ANOVA revealed a significant effect of block on performance $F(5, 95) = 10.058$, $p < 0.001$; there was also a significant linear trend $F(1, 19) = 26.145$, $p < 0.001$.

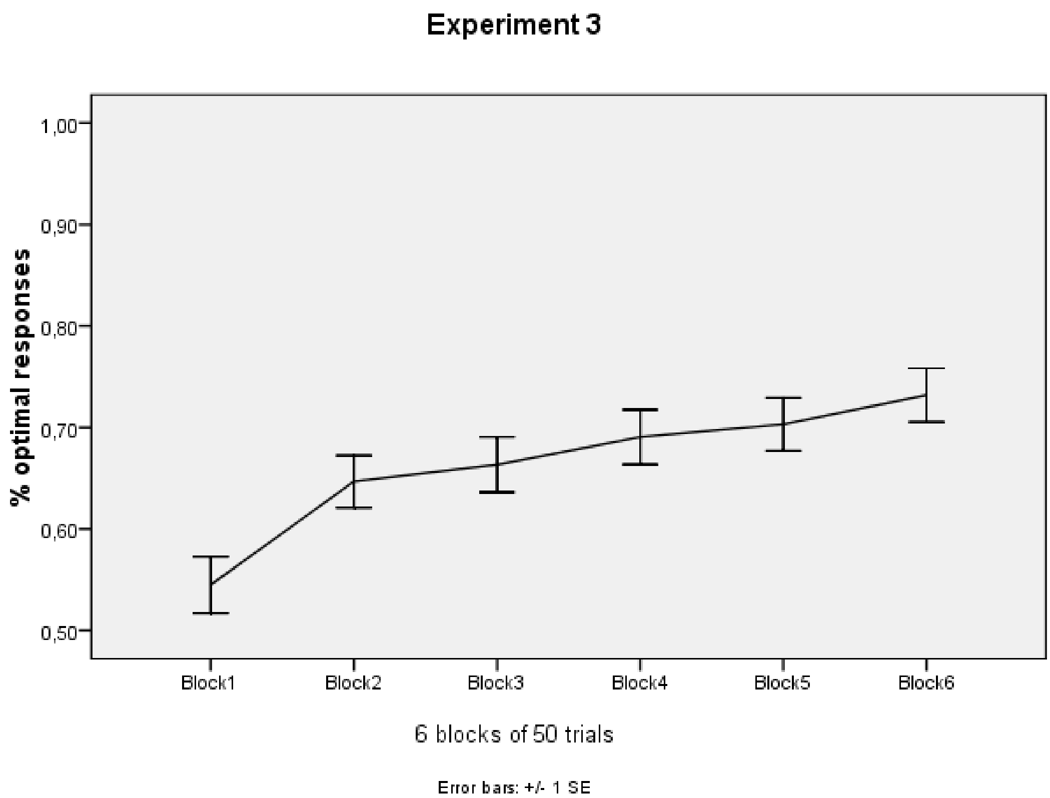

Figure 2.10. Experiment 3 behavioural data from the categorization task, analysed in blocks of 50 trials.

2.3.2.2. Test Task

The main finding from the post-training test task was participants' reluctance to write down their replies. Only 2 out of 20 participants filled in the questionnaire without complaints, while the rest of them expressed their great difficulty in completing the required task, typically saying “How can I write down the sounds?”, while reporting that they “had in their minds” the sounds probed for. These reluctant participants were gently encouraged to fill in the questionnaire, and to use the first word that came in their minds to describe the sounds.

The names participants gave for the sounds were characterized by a great variability. We used a success criterion of 65% optimal responses in the final block of 50 trials, and we examined the replies of 16 participants (out of 20) that met this criterion. We were unable to distinguish the tones by the names they provided (e.g. it was not clear which tone they referred to when they said “melodic sound”), so their replies could not be scored as in Experiment 2 and no quantification of the data was extracted. Instead, supposing that the two names they provided were indeed meant to describe the cues that were highly predictive of sun and rain respectively, the 32 names referring to the four cues are shown in Table 4.

Table 4

Names That Were Given for the Four Complex Tones by Participants in the Test Task, Experiment 3.

Cues	Names
1	Intense, Thick, Heavy (2), Melodic, Deep, Short (2)
2	Thin (5), Subdued, Shrill, High, Less abrupt
3	Shrill, Thin (3), Heavy, Deep, Bass, Brief, Low, Sharp
4	Heavy (2), High, Deep, Abrupt

Note—numbers in parenthesis, if present, denote the number of participants that gave the same answer. The names shown were translated from Greek.

2.3.3. Discussion

Results from Experiment 3 indicate that participants improved in their ability to categorize the combinations of complex tones as the training progressed. Although we considered the task quite difficult to accomplish, participants were, to our surprise, able to solve the task. Yet, their performance over 302 trials (66.28%) was considerably lower than the performance of participants over 200 trials in Experiment 2 (75.24%).

Moreover, the reluctance observed in answering the questionnaires and the variability in the names participants provided for the complex tones indicate that the complex tones were difficult to name, or, at least, that there are not unanimous, easily retrievable names for the tones we utilized as cues in Experiment 3. Thus, the characterization of this condition as “non-nameable” condition, we consider, has been experimentally justified.

2.4. Between subjects analysis

In order to test our prediction (as described in section 1.3.3) a between-subjects analysis was conducted. In Fig. 2.11 categorization performance of the three groups participating in each of the experiments is shown (for Experiment 2 data from the 2nd group are depicted). The comparison of interest is that between performance of participants in Experiment 2 (“nameable” condition) and performance of participants in Experiment 3 (“non-nameable” condition). As evident in Fig. 2.11, there is a difference in performance between the two groups, with participants in the nameable condition outperforming participants in the non-nameable condition throughout training. Analysis, though, in blocks of 50 trials masks performance levels at the beginning of training, and does not provide the opportunity to check our prediction. Thus, we analysed categorization data from Experiments 2 and 3 in blocks of 10 trials and the corresponding learning curves are shown in Fig. 2.12.

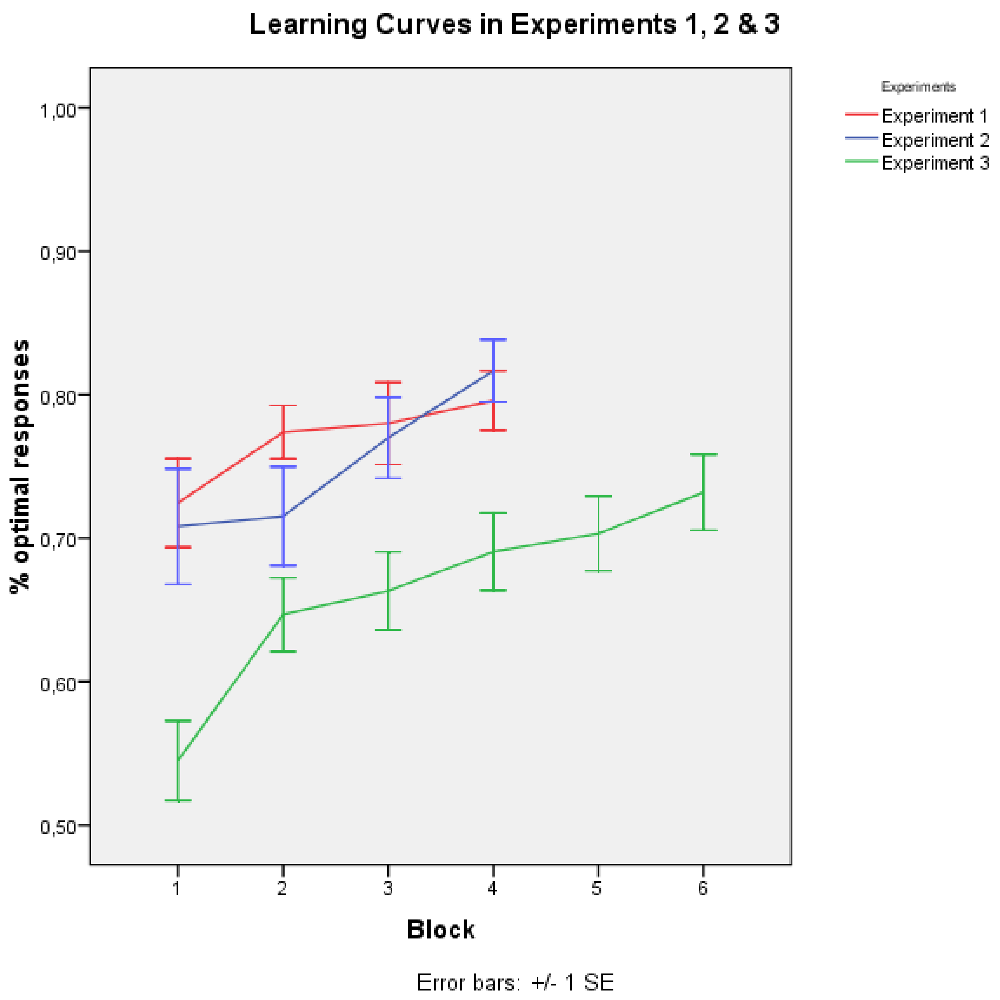

Figure 2.11. Categorization performance of participants in Experiments 1, 2, and 3, analyzed in blocks of 50 trials.

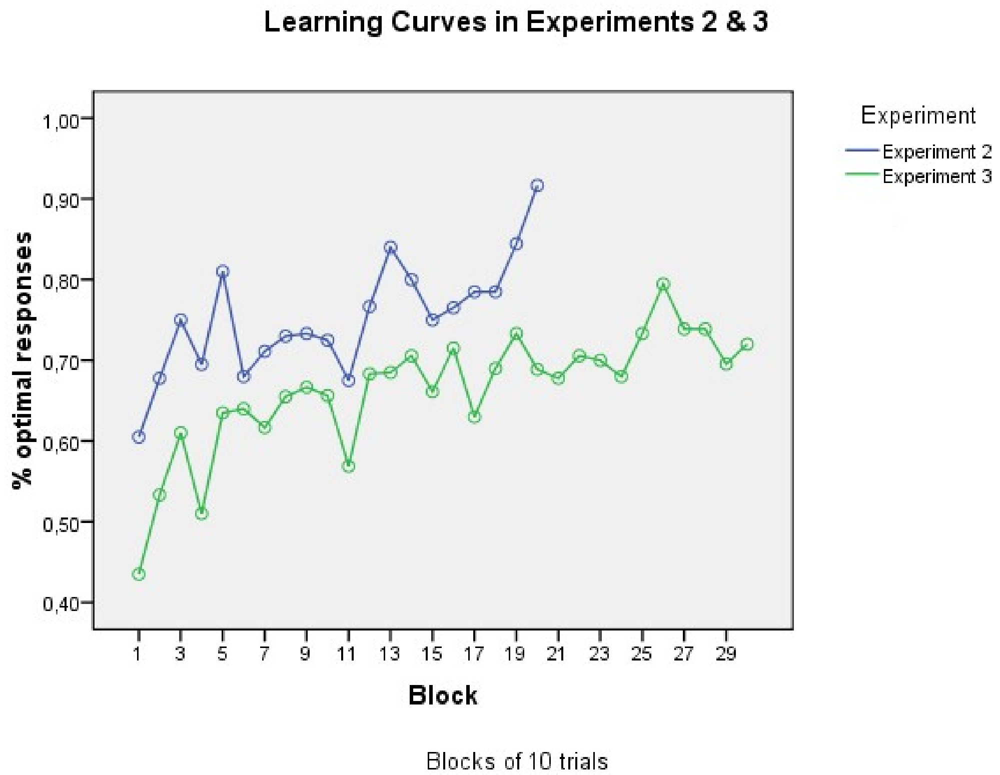

Figure 2.12. Categorization performance of participants in Experiments 2 (nameable condition) and 3 (non-nameable condition), analyzed in blocks of 10 trials.

As can be seen in Fig. 2.12, participants in the condition with the nameable stimuli outperformed participants in the condition with the non-nameable stimuli from the very start of the training by approximately 10%. This difference in performance was maintained throughout training. We conducted a between-subjects ANOVA of the data from the first 200 trials (that were available by both groups) with trial block as the within-subjects factor and condition as the between-subjects factor. The interaction of block by condition was found not to be statistically significant, $F(19, 722) = 0.93$, $p = 0.544$, while the main effect of block on performance was revealed to be significant, $F(19, 722) = 6.93$, $p < 0.001$ (Greenhouse-Geisser corrections are not reported, as in section 2.2.2.1.), which indicates that subjects improved in their categorization ability no matter the condition they participated in. The main effect of condition on performance was also found to be significant, $F(1, 38) = 13.056$, $p = 0.001$, indicating that performance in the nameable condition was higher compared to the one in the non-nameable condition, independently of the stage in training.

The same analysis was performed on categorization data combined from the 1st group participating in Experiment 2 and the group that participated in Experiment 3. We analyzed data in blocks of 10 trials and the results were statistically the same, namely the interaction of block by condition wasn't significant, $F(19, 722) = 1.295$, $p = 0.178$ whereas there was a main effect of block on performance, $F(19, 722) = 5.701$, $p < 0.001$ and also a main effect of condition on performance $F(1, 38) = 8.934$, $p = 0.005$ (Greenhouse-Geisser corrections on degrees of freedom not reported).

In order to more accurately check our prediction, we needed to examine if performance in the non-nameable condition could be considered to be at chance levels for a longer period in training compared to performance in the non-nameable condition. We used confidence intervals (at the level of 95%) of the mean performance per condition and per block of 10 trials, and we examined if chance performance, namely 50% (Knowlton et al., 1994), was included in these intervals. Performance levels of the groups participating in each condition in blocks of 10 trials with error bars being confidence intervals are depicted in Fig. 2.13.

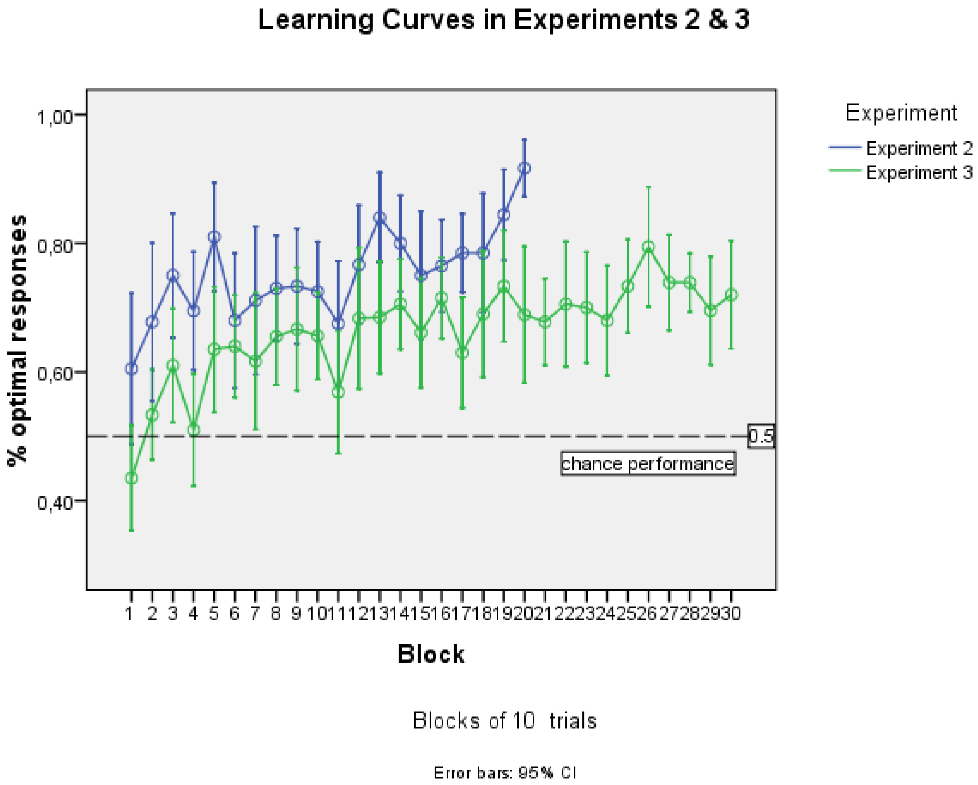

Figure 2.13. Categorization performance of participants in Experiments 2 (nameable condition) and 3 (non-nameable condition), analyzed in blocks of 10 trials, along with confidence intervals and chance performance.

For clarity reasons, in Fig. 2.14 we depict only the first three blocks of 10 trials. It can be seen that performance of participants in Experiment 2 (nameable condition) can be considered to be at chance levels only for the first block of 10 trials, whereas the same holds for both the first and the second block of 10 trials for performance of participants in Experiment 3 (non-nameable condition).

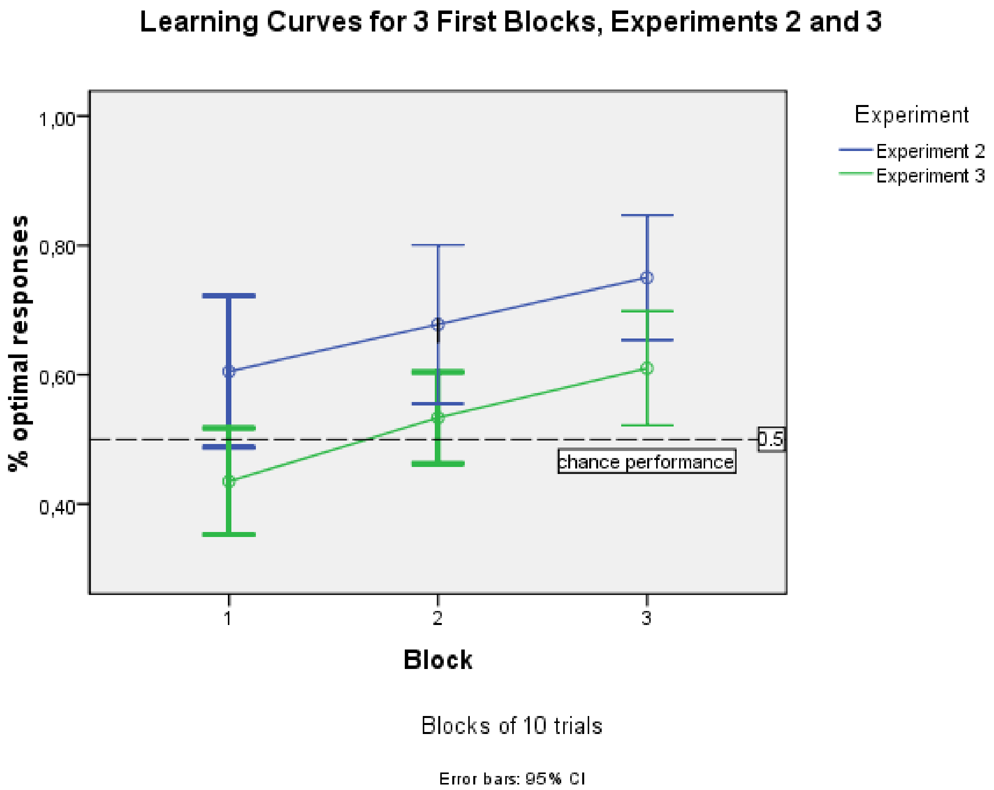

Figure 2.14. Categorization performance of the first three blocks of 10 trials of participants in Experiments 2 (nameable condition) and 3 (non-nameable condition).

As in the case of the mixed ANOVA reported previously, the same analysis was performed with data on Experiment 2 taken from the first group of participants. As can be seen in Fig. 2.15 the results are repeated, namely performance of participants in non-nameable condition can be considered to be at chance levels for a longer period initially in training (first and second trials blocks), compared to performance of participants in the nameable condition (first trial block). Thus it has been proven reasonable to have assumed that the discrepancy in performance at later stages of training between the two groups participating in Experiment 2 has no effect on the verification of our prediction.

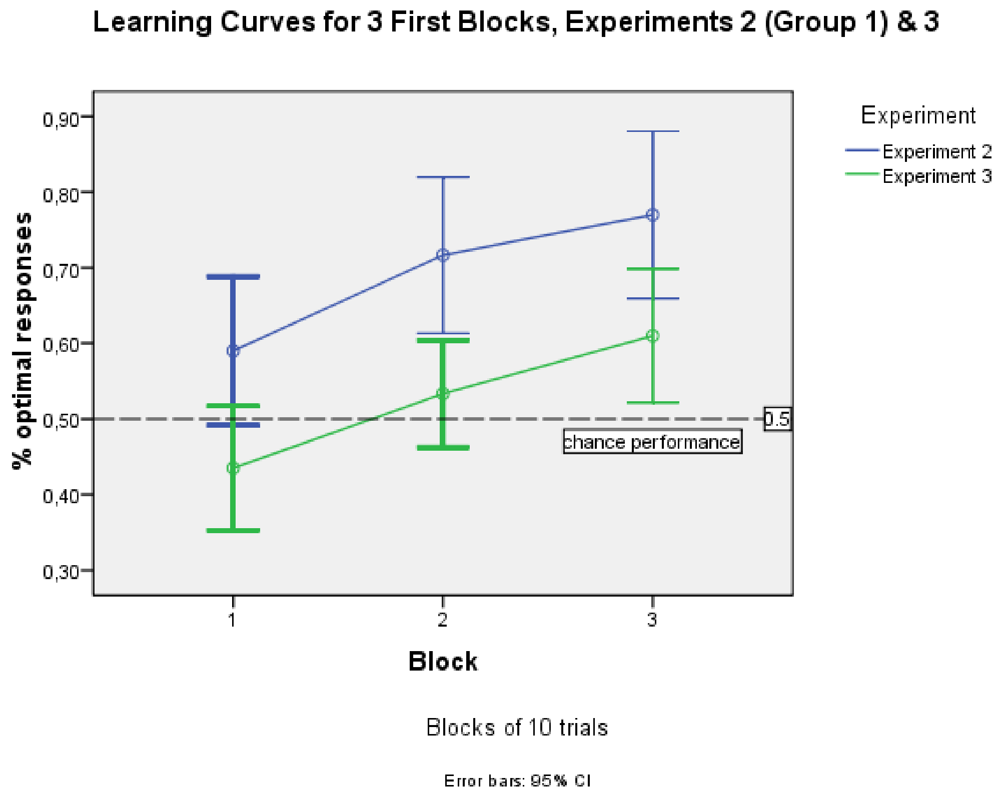

Figure 2.15. Categorization performance of the first three blocks of 10 trials of the first group of participants in Experiment 2 (nameable condition) and participants in Experiment 3 (non-nameable condition).

In order to investigate whether participants in the nameable and in the non-nameable condition exhibited the same or different learning rate during the categorization task we conducted linear regression analysis on average performance data, analyzed in blocks of 10 trials, from participants in Experiment 2 (2nd group) and Experiment 3. In our analysis we excluded the near-chance initial periods in performance for Experiment 2 (first block) and Experiment 3 (first and second blocks).

In the nameable condition it was found that B = 0.008, 95% C.I. [0.004, 0.012] whereas in the non-nameable condition B = 0.005, 95% C.I. [0.003, 0.007]. Since confidence intervals overlap, it could be argued that participants in the non-nameable condition exhibited during the task comparable learning rate with participants in the nameable condition. Therefore, it could also be argued that with further training participants in the non-nameable condition (achieving on average in block 30 73.19%

optimal responses) could reach by the end of the task performance levels of participants in the nameable condition (achieving in block 20 81.66% optimal responses). Clearly this assumption requires further empirical investigation.

2.4.1. Discussion

Results from the between-subjects analyses revealed that participants in Experiment 2 (nameable condition) had an advantage in the categorization task from the very start of it relative to participants in Experiment 3 (non-nameable condition). This advantage was maintained throughout training, as participants in the nameable condition constantly outperformed participants in the non-nameable condition by approximately 10%.

Also, confidence intervals of the mean performance of participants in Experiments 2 and 3 provided evidence indicating that performance of participants in the non-nameable condition can be considered to be at chance levels for a longer period early in training compared to performance of the groups that took part in the nameable condition. This result seems to verify our prediction.

Part 3

General Discussion

One significant finding of the present research is evidence indicating gradual learning in an auditory version of the WPT in which the cues are sequentially presented to participants. In Experiments 2 and 3 participants seem to have acquired knowledge of the associations between the four auditory cues and the two outcomes, evident in the fact that performance gradually increased throughout training. This result, to the best of our knowledge, has not been reported before.

Our main finding, though, is that participants in the nameable condition (Experiment 2, in which cues were animal sounds) constantly outperformed participants in the non-nameable condition (Experiment 3, in which cues were computer-generated complex tones) by approximately 10% throughout the training period. We experimentally verified that the animal sounds were easy to name, as opposed to the complex tones that participants found difficult to name.

We hypothesized that verbalizability is an operating characteristic of declarative knowledge and therefore assumed that participants in the non-nameable condition would be unable to develop, or rely on, declarative verbalizable rules early in training. We thus predicted, based on previous reports of healthy participants' performance in the WPT, that performance of participants in the non-nameable condition would be at chance levels for a longer period initially in training, compared to performance of participants in the nameable condition.

Indeed such a between-subjects difference in performance was revealed, and we attribute this difference to the names available for the animal sounds to participants in the nameable condition, as opposed to non-availability of such names for the complex tones to participants in the non-nameable condition.

3.1. Open questions

We assumed that participants in the non-nameable condition might develop arbitrary names for the cues of the task at some point during training, and therefore be able, from that point onwards, to also develop and rely on verbalizable rules. It could though be argued that participants never developed such names during the task. The requirements of the test task used to assess verbal behaviour of participants could have forced participants to declare arbitrary names for the cues at that point, when probed for.

Our experimental design cannot empirically distinguish between the two options. In the case that participants developed arbitrary names for the cues, performance later in training could be guided by verbalizable rules, mediated by the declarative memory system. In the case that no names had been developed during the task the increase in participants' performance in the non-nameable condition could be supported by implicit, procedurally mediated knowledge of the cue-outcome associations. Each assumption is equally compatible with our prediction since the difference in performance predicted is with respect to the first few trials in training.

We do believe though, that, as a result of our experimental manipulation, participants in the non-nameable condition were discouraged in using verbalizable rules throughout the training, and hence the contribution of the declarative memory system in this condition was overall decreased relative to the nameable condition. If this assumption is correct, then our behavioural findings could be interpreted as follows: Participants who could develop and rely on verbalizable rules mediated by the declarative memory system constantly outperformed participants who were unable to develop or rely on such rules. Support to this notion is offered by reports on hypoxic and control participants' performance by Hopkins et al. (2004). Based on the fact that hypoxic patients are characterized by a deficit in the acquisition of new declarative memories it could be argued that such patients are unable to develop, or maintain verbalizable rules mediated by the declarative memory system. Thus, the Hopkins et al. findings could be interpreted in a similar manner. Healthy participants who could develop and rely on verbalizable rules constantly outperformed hypoxic patients who were unable to develop or maintain such rules. Moreover, according to Foerde et al. (2006), making the cue-outcome associations less probabilistic enhances the declarative learning of cue-outcome associations. Such a manipulation was conducted be Gluck et al. (2002) and the findings can be likewise interpreted. In their Experiment 2 the “newer” version of the task was used (thought to enhance declarative learning) and participants' performance was constantly greater than performance of participants in their Experiment 1 in which the cue-outcome association were less probabilistic. In addition, a similar account was provided by Foerde et al. (2007). Analysis of participants' performance in the ice cream task under dual-task versus single-task conditions and of corresponding measures of explicit knowledge of the cue-outcome associations showed that “the subjects with at least some awareness of the cue-outcome associations performed better than those with no apparent awareness” (p. 873). Since awareness is thought to be supported by the declarative memory system (Reber et al., 1996), it could be argued that participants who were able to acquire declarative knowledge of the task performed better than those who were not. Finally, the most striking support to our understanding of the difference in performance between the nameable and the non-nameable condition is offered by the study of Price (2009). Price reduced the time made available to healthy participants to process the feedback after their response, a manipulation thought to disrupt explicit (or declarative, according to Reber & Squire, 1994) hypothesis-testing processes. In her Experiment 2 a group of participants were given a 2.5 sec delay after the delivery of feedback in a WPT trial, before performing a memory scanning task, (Long Feedback condition) whereas another group of participants were required to perform the memory scanning task immediately after the delivery of feedback (Short Feedback condition). Price reported that her “manipulation, designed to disrupt explicit-learning processes, substantially reduced classification accuracy” (p. 210) and our interpretation of her findings is quite similar. As can be seen in Fig. 3.1, participants thought to be able to engage in declarative explicit-learning processes (Long Feedback condition) constantly outperformed participants that were thought to be unable to engage in such declarative processes (Short Feedback condition). Overall, all of the above interpretations seem to favour our belief that our Experiment 3, utilizing cues that are difficult to name, provides a version of the WPT that discourages the use of declarative strategies, and therefore provides an even better way of assessing the gradual acquisition of cognitive skills according to Knowlton et al. (1994).

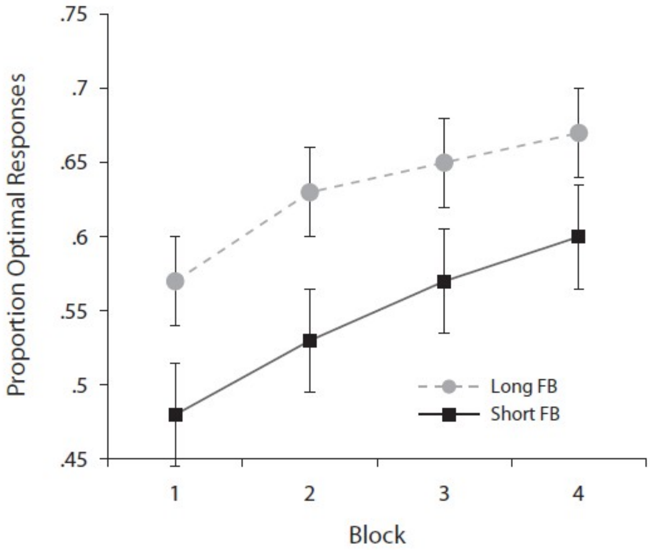

Figure 3.1. Categorization performance of healthy participants in the long- and short-feedback conditions in the WPT, Experiment 2, Price, 2009 (p. 217).

The aforementioned interpretation of the Hopkins et al. (2004) findings also offers an intriguing prediction. According to Knowlton et al. (1994) amnesic patients are impaired in forming declarative memories able to support performance, and therefore, we assume, rely on implicit, procedurally mediated knowledge in solving the WPT. Thus, we predict, their performance in the nameable condition will not differ from performance in the non-nameable condition since the difference between the two conditions is the ability of participants to develop or rely on declarative, verbalizable rules. A neuropsychological study examining this prediction is likely to provide interesting findings.

Another way of examining our hypothesis that using non-nameable cues in the WPT discourages the development or use of declarative strategies is by investigating brain activity of healthy participants in the nameable and the non-nameable condition. A neuroimaging study, assessing the differential activation of the brain structures thought to underlie the declarative and the procedural memory system during the expression of knowledge acquired in the two conditions (in probe trials after the categorization task, following Foerde et al., 2006) might shed light on the issue. Our prediction is that MTL will be more activated during probe trials of items learned in the nameable condition relative to items learned in the non-nameable condition.

The previous interpretations and predictions are based on the assumption that participants in the non-nameable condition never developed arbitrary names for the cues, and were therefore unable to develop, or rely on, verbalizable rules. Although we believe this notion to be true for the majority of participants in Experiment 3, we cannot ignore alternative accounts. In a study by Galizio & Baron (1976) it was suggested that changes in participants' performance during an auditory categorization task might “reflect (verbal) labels subjects assign to the stimuli” (p. 592). Therefore, it could be argued that some of the participants in the non-nameable condition developed arbitrary names for the cues, whereas others did not. As mentioned in Section 2.3.2.2, two out of twenty participants in Experiment 3 responded immediately to the questionnaire administered to them. Also, one more participant (data were not included in the analysis due to technical problems) when asked to comment retrospectively on her strategy during the task stated that “Until almost the first break in the experiment, I was trying to name the sounds” (it is worth recalling that none of the participants had ever been informed of our hypothesis). Perhaps an analysis of participants' reaction times in responding to the test task could distinguish between those participants that presumably used verbalizable rules (rapidly naming the cues) and those who didn't (reluctant in naming the cues). In a neuroimaging experiment with many more participants, more of them might be fast in providing names for the cues and thus a statistical analysis might be possible. A correlation analysis between participants' reaction times in the test task and task-related activity in the MTL and basal ganglia during the first few trials in training or during expression of knowledge might provide more evidence in favour of our assumption that verbalizability is an operating characteristic of declarative knowledge. We predict that activation of the MTL will be greater for participants immediately naming the cues relative to reluctant participants.

Another issue which has implications for the contribution of the declarative and the procedural memory system to the acquisition of knowledge in the WPT is the form of feedback used in our implementation. It has been suggested that basal ganglia are involved in reward prediction, thus the form of feedback and the timing of feedback utilized in the WPT seem to have an effect on healthy individuals and patients' performance (Shohamy et al., 2008).

As noted previously, there is a discrepancy in the form of feedback provided to participants between the typical implementation of the WPT and our auditory version of the task. Participants in the prototypical WPT are presented, after their prediction, with the actual outcome, while the combination of cues stays on the screen. On the other hand (see Section 2.2.1.2.), participants in our auditory version are presented with the actual outcome while the combination of auditory cues is not repeated. The former form of feedback is thought to encourage the engagement of the declarative memory system, as assumed for the paired associate task of Experiment 1 in the Poldrack et al. (2001) study. Thus, the latter form of feedback, we assume, does not encourage the engagement of the declarative memory system during the categorization task. We also assume that the form of feedback implemented in our auditory version implicates more working memory resources since participants would have to keep in their working memory the combination of cues in order to associate them with the actual outcome. According to Foerde et al. (2006), declarative memory performance is thought to depend on working memory resources, but the effect of working memory demands on the engagement of the declarative and the procedural memory systems in the WPT has only been examined with respect to a secondary task (Foerde et al., 2006, 2007; Newell et al., 2007). So, to our knowledge, this purported difference in working memory resources does not provide any indications as to the engagement of each memory system in the WPT. Overall, we assume that the form of feedback utilized in our auditory version of the WPT might render the task less declaratively mediated, since it does not encourage declarative encodings of the stimuli. Our experimental design cannot isolate the effect of the form of feedback on the engagement of the procedural and the declarative memory system in the WPT, since the prototypical feedback was utilized in a visual version of the WPT, while our form of feedback was utilized in auditory versions. An experiment assessing the effect of the form of feedback (utilizing the two forms in two visual versions of the WPT) on brain activity during training or during the expression of knowledge might support evidence in favour of our assumption that our form of feedback can render learning in the WPT less declaratively mediated.

3.2. Limitations of the present study

The main finding of our research, namely that participants in the nameable condition of our experiment constantly outperformed participants in the non-nameable condition, was attributed, according to our understanding, to the fact that the animal sounds in the former condition were easy to name, as opposed to the fact that the complex tones were difficult to name.

It could be argued though, that nameability of the stimuli in Experiment 2 as opposed to non-nameability of the stimuli in Experiment 3 is not the sole difference between the two conditions. Indeed, it has been suggested to us (Prof. K. Moutoussis, personal communication) that other factors, such as different duration, perceptual intensity, or discriminability of the stimuli that served as cues in the two conditions might be responsible for the observed difference in performance between the two conditions.

One counter-argument against the effect of stimuli's perceptual features on performance is offered by Knowlton et al. (1994). In that study, three different tasks were implemented and in each task three different kinds of stimuli were utilized (see Fig. 1.2). As mentioned previously (Section 1.1.2.), a three-way ANOVA (group x task x trial block) on combined data from healthy participants and amnesic patients' performance in the first 50 trials revealed no interaction of the factors and no significant main effect of task on performance. Since stimuli in the three tasks were highly dissimilar but all other procedural details were identical, we assumed that this finding suggests a lack of effect of the stimuli's perceptual features on early performance.

But indeed, our results indisputably indicate an effect of the difference in the characteristics of the cues (being nameability or some other perceptual factor) on performance, weakening our counter argument. Moreover, it cannot be firmly assumed that this lack of effect is bound to be repeated in an auditory version of the WPT as ours, neither that it is maintained in later performance. Clearly an equation of the stimuli's perceptual features between the two conditions is required in order to support that the observed finding can be attributed to difference in the nameability of the cues used in the two conditions.

A suggested manipulation (Prof. G. Gyftodimos, personal communication) was to implement a newer version of the nameable condition by using the same complex tones as in the non-nameable condition, but to previously train participants in naming the cues by means of a repeated naming task. Once it is experimentally verified that participants indeed have acquired unique and consistent names for the four complex tones then they should be administered the categorization task. Yet, research on label training of auditory stimuli has suggested that such training improves subsequent discriminability of stimuli (Galizio & Baron, 1976). Also, learning to associate sounds with verbal labels given by the experimenter has been found to improve participants' subsequent ability to identify the sounds compared to a condition in which the sounds were associated with graphic images that were also given by the experimenter (Edworthy & Hards, 1999). In order to equate for an expected increase of the cues' discriminability due to label training, a control learning task with similar attentional demands (as suggested by Galizio & Baron, 1976), identical exposure to the cues, and associational requirements leading to same post-training memory performance should be administered to participants in a newer version of the non-nameable condition. Moreover, this control learning task should not encourage participants to name the stimuli of the task. A potential difference in participants' performance between the two newer versions of the nameable and the non-nameable condition can be then more firmly attributed to the difference in nameability of the cues between the two conditions. Indeed, such a control experiment is in our immediate plans, and we hope that it will provide evidence in favour of our assumption that verbalizability is an operating characteristic of declarative knowledge.

3.3. The order of trials in the WPT

As noted previously, the order of trials utilized in the WPT is random but the same for all participants in the task. This order seems to be a part of the WPT paradigm, followed by most researchers that investigated the cognitive mechanisms of either healthy (e.g., Gluck et al., 2002; Meeter et al., 2006, 2008) or clinical populations (e.g., Hopkins et al., 2004; Shohamy et al., 2004).

The effect of the order of trials and of the corresponding actual outcome on participants' behaviour has not received wide attention, with—to the best of our knowledge—one exception. Foerde et al. (2007) examined the effect of a secondary task on participants' performance in the ice cream task and commented that “the ST group was able to take advantage of an early trial repetition” (p. 867). This comment seems to imply that the distribution of patterns across trials has an effect on performance.

Moreover, despite the fact that the WPT has been used to assess the gradual learning of cognitive skills, recent accounts suggest that this notion might not be true. Meeter et al. (2008) reported a pattern of non-linear, or “jerky”, performance by participants. The authors applied strategy analysis (Meeter et al., 2006) and rolling regression analysis (Lagnado et al., 2006) on behavioural data by Gluck et al. (2002).

Their results showed sudden shifts in participants' behaviour, which seem to indicate that participants during the task “shift from one rule to another” (Meeter et al., 2008, p. 239). A similar report on a change in participants' behaviour throughout the task comes from a note by Hopkins et al. (2004) on participants' frustration.

Clearly, thus, there is a change in participants' behaviour throughout the task, evident in their non-linear performance (Meeter et al., 2006, 2008) and in their frustration (Hopkins et al., 2004). In elucidating the cognitive mechanisms underlying performance in the WPT it seems important to determine the factors that affect behaviour during the task.

Since every other procedural detail remains unaltered throughout the trials of the task, it seems reasonable to suggest that the cause of participants' frustration (and possibly of their non-linear performance) are the events in which the less probable outcome on a pattern appears (through feedback) to be the correct one. By inspecting the standard order of trials typically utilized in the WPT, we noticed that such events do not occur early in the task (the first less probable outcome appears in trial 23 as can be seen in the Appendix), rendering the task deterministic, at least early in training, and, perhaps, enabling participants to rapidly achieve better-than-chance performance. This assumption was strengthened by our experience with 80 participants in the WPT. We noticed that most of our participants got frustrated when the actual outcome on a pattern (especially on a one-cue pattern) was the less probable one. We therefore assume that the event of a less probable outcome falsifies a participant's so-far accumulated knowledge and we argue that those events are the cause of participants' change in behaviour.

It seems reasonable to suggest, though, that not all patterns in the task are characterized by the same “predictive value”, that is, not all patterns contribute equally to learning, at least if explicit knowledge supports performance. Foerde et al. (2007) reported—when assessing participants' explicit knowledge of the task—that “the subjects were better at selecting the figure (pattern) associated with a given outcome when only one cue was present” (p. 868). This report suggests that the number of cues comprising a pattern might have an effect on the pattern's predictive value. Also, in the analysis of performance in the WPT, as introduced by Knowlton et al. (1994), participants' responses on two of the patterns in the task (F and I) that predict the two outcomes equally often are excluded from the analysis. This fact also suggests that these two patterns do not contribute to learning and we argue that the predictive value of each pattern is related to the probability of an outcome given that the pattern is presented.

We suggest that the effect on participants' behaviour of a less probable outcome on a pattern is inversely related to the pattern's predictive value. Thus, we propose a quantification of the “inconsistency” caused during the course of learning by an event of a less probable outcome on a trial in the WPT, which should be related the two aforementioned factors, namely by the number of cues in a presented pattern and by the P(outcome|pattern) quantity.

We suggest an implementation of various versions of fixed order of trials in a random manner. Each version should be administered to a group of participants. By summing the inconsistency of each trial in a small window preceding a participant's response, we can create a quantification of the inconsistency of this window. We predict that the points of sudden jumps in behaviour (as can be revealed by the methods of Meeter et al., 2008) or simply participants' performance will be correlated with the window's inconsistency. If this prediction is empirically verified, then it will provide evidence in favour of a hypothesis-testing cognitive process during acquisition of knowledge, which in turn is consistent with the assumption of rule-based learning[1] in the WPT (Meeter et al., 2006).

Moreover, we suggest the implementation of a semi-random order of trials, devised in a way so that the measure of inconsistency of trial-windows will be characterized by a uniform distribution throughout the task. We predict that if this version is administered to a group of healthy individuals then the analysis of their performance will not reveal points of shift in their behaviour, or, at least, that the points of shift will be either less apparent or less in number compared to the ones in a group administered the WPT with the standard order of trials. Such a finding would be consistent with the assumption of an incremental learning cognitive mechanism supporting performance in the WPT (assumed by Lagnado et al., 2006, according to Meeter et al., 2008), and moreover, it would suggest that the involvement of either incremental or rule-based learning mechanisms during acquisition of knowledge in the WPT might be experimentally manipulated.

[1] Rule-based learning does not imply that the acquired knowledge is procedurally mediated as the definition of procedural knowledge as rule-based by Cohen & Squire (1980) would suggest. According to Meeter et al. (2008) “Rule-based learning could be seen as either declarative, as the rules are presumably flexible and representational, or procedural, as the rules describe how to do a task” (p. 238).

3.4. Hippocampal Representations

One of the most important open questions in the study of probabilistic category learning is "what is learned" (Shohamy et al., 2008). The nature of the internal representations thought to support performance in the task, mediated by either MTL or basal ganglia, has not yet been clarified.

Gluck and Myers (1993) proposed a computational theory of the hippocampal region's function. According to this theory the hippocampal region develops new stimulus representations by altering the similarities among pre-existing representations. The new representations are formed according to two basic principles, "predictive differentiation" and "redundancy compression", which—in summary—map together stimuli predicting the same outcome or distinguish stimuli predicting alternative outcomes. This theory was applied to data from Knowlton et al. (1994) by means of a connectionist model of category learning (Gluck et al., 1996). According to this adaptation, new representations that are formed by the hippocampal module may eventually be adopted by other brain structures such as the prefrontal cortex or basal ganglia.

The Gluck & Myers (1993) computational theory has been repeatedly linked to neuroimaging findings (Poldrack et al., 2001) reporting an activation of the MTL early in training and a deactivation of this brain region in later phases of the task (eg., Shohamy et al., 2008; Meeter et al., 2008). According to this prevailing interpretation, the medial temporal lobe sets up new representations early in the task that are then used by other regions such as the striatum to develop complex stimulus-response associations (Poldrack et al., 2001) that allow higher levels of accuracy and faster learning later in the task (Shohamy et al., 2008).

Perhaps a way of assessing the nature of MTL-mediated representations is by assessing the nature of those pre-existing representations that the hippocampal region is assumed to re-encode. Shohamy et al. (2008) suggested that the MTL supports the formation of representations based on relations between the stimuli, such as "red beats yellow" (p. 227) or perhaps "squares beats circles", applying this notion to the WPT paradigm. Although Shohamy et al. never implied that the pre-existing representations are linguistic ones, it could be argued, given that "humans are the only animals to associate conceptual categories with words" (Lupyan, 2006), that in human cognition the MTL-mediated representations formed during a categorization task are re-encodings of perceptual pre-existing representations associated with verbal labels.

We suggest that the findings (if replicated, and if not falsified by the control experiments described previously) and the experimental manipulation of the present study might provide a means of assessing the hypothesis that MTL-mediated re-encodings are based on linguistic pre-existing representations.

If our hypothesis is correct then participants who have not been trained in naming auditory tones serving as cues in the WPT would be unable to form new representations. To the contrary, participants who have acquired consistent names for the tones through a repeated naming task, would be able to form new, MTL-mediated, representations. The activation of MTL in participants in the former case, we predict, will be decreased relative to the activation of MTL in participants in the latter case, at least early in training, presumably for the first 15 trials, as suggested by Meeter et al. (2008). A neuroimaging study examining the aforementioned predictions might shed light on the nature of declaratively re-encoded representations.

3.5. Conclusions

In summary, we manipulated participants' ability to name the cues in an auditory version on the WPT and we argued that the observed differences in performance between a group participating in a condition utilizing recognizable, easy-to-name animal sounds and a group participating in a condition utilizing difficult-to-name complex tones supports the hypothesis that verbalizability is an operating characteristic of declarative knowledge. The experimental manipulation and the findings of the present study offer intriguing predictions as to the contribution of the declarative and the procedural memory to the acquisition of knowledge in the WPT by healthy and clinical populations, and also have implications as to the nature of MTL-mediated re-encodings.

References

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, *51*, 355–365.

Anderson, J. R. (2007) *How Can the Human Mind Occur in the Physical Universe?* New York: Oxford University Press.

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

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *210*, 207–209.

Dienes, Z., & Perner J. (1999). A theory of implicit and explicit knowledge. *Behavioural and Brain Sciences, 22*, 735-808.

Edworthy, J., & Hards, R. (1999). Learning auditory warnings: the effect of sound type, verbal labelling and imagery on the identification of alarm sounds. *International Journal of Industrial Ergonomics, 24,* 603-618.

Egeth, H. E., & Yantis S. (1997). Visual Attention: control, representation, and time course. *Annual Review of Psychology*, *48*, 269-297.

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Science, U S A*, *103*, 11778-11783.

Foerde, K., Poldrack, R. A., & Knowlton, B. J. (2007). Secondary task effects on classification learning. *Memory & Cognition, 35* (5), 864–874.

Foerde, K., & Poldrack, R. A. (2009). Procedural Learning in Humans. In Squire, L.R., (Ed.) *The New Encyclopedia of Neuroscience*, *7*, 1083-1091.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers*, *35*, 116–124.

Galizio, M., & Baron, A. (1976). Label training and auditory categorization. *Learning and Motivation, 7,* 591-602.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: an adaptive network model. *Journal of Experimental Psychology: General, 117*, 227–247.

Gluck, M. A., & Myers, C. E. (1993). Hippocampal mediation of stimulus representation: a computational theory. *Hippocampus*, *3*, 491–516.

Gluck, M. A., Oliver, L. M., & Myers, C. E. (1996). Late-training amnesic deficits in probabilistic category learning: A neurocomputational analysis. *Learning & Memory, 3*, 326–240.

Gluck, M. A., Shohamy, D., & Myers, C. E. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic classification learning. *Learning & Memory*, *9*, 408–418.

Hardin, C. L. (2005). Explaining basic color categories. Cross-Cultural Research, *39* (1), 72-87

Holt, L., & Lotto A. J. (2006). Cue weighting in auditory categorization: Implication for first and second language acquisition. *J. Acoust. Soc. Am., 119* (5), 3059-3071.

Hopkins, R. O., Myers, C. E., Shohamy, D., Grossman, S., & Gluck, M. A. (2004). Impaired probabilistic category learning in hypoxic subjects with hippocampal damage. *Neuropsychologia, 42*, 524–535.

Kay, P., & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Science, U S A*, *100*, 9085–9089.

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.

Levitin, D. J., & Rogers, S.E. (2005). Absolute pitch: perception, coding and controversies. *Trends in Cognitive Sciences, 19* (1), 26-33.

Lewis, J. W., Wightman, F. L., Brefczynski, J. A., Phinney, R. E., Binder, J. R., & DeYoe, E. A. (2004). Human brain regions involved in recognizing environmental sounds. *Cerebral Cortex, 14* (9), 1008-1021.

Lupyan, G. (2006). Labels facilitate learning of novel categories. In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The Sixth International Conference on the Evolution of Language* (p. 190-197). Singapore: World Scientific.

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

Meeter M., Radics G., Myers C. E., Gluck M. A., & Hopkins R. O. (2008). Probabilistic categorization: how do normal participants and amnesic patients do it? *Neuroscience and Biobehavioural Reviews, 32*, 237-248.

Moody, T. D., Bookheimer, S. Y., Vanek, Z., Knowlton, B. J., 2004. An implicit learning task activates medial temporal lobe in patients with Parkinson's disease. *Behavioral Neuroscience*, *118*, 438-442.

Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2007). Challenging the role of implicit processes in probabilistic category learning. *Phychonomic Bulletin & Review, 14* (3), 505-511.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1–32.

Pinker, S., & Jackendoff, R. (2005). What's special about the human language faculty? *Cognition, 95*, 201-236.

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

Poldrack, R. A., & Rodriguez, P. (2004). How do memory systems interact? Evidence from human classification learning. *Neurobiology of Learning and Memory, 82*, 324–332.

Poldrack, R. A., & Foerde K. (2008). Category learning and the memory system debate. *Neuroscience and Behavioural Reviews, 32*, 197-205.

Price, A. L. (2009). Distinguishing the contributions of implicit and explicit processes to performance of the weather prediction task. *Memory & Cognition, 37*, 210-222.

Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39* (4), 859–862.

Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning & Memory*, *1*, 217–229.

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

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning Memory and Cognition*, *20*, 585–594.

Schacter, D. L. (1987) Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 501–18.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neuropsychiatry and Clinical Neuroscience*, *20*, 11–21.

Seitz, A.R., Protopapas, A., Tsushima, Y., Vlahou, E.L., Gori, S., Grossberg, S., & Watanabe, T. (2010). Unattended exposure to components of speech sounds yields same benefits as explicit auditory training. *Cognition, 115*, 432-443.

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? *Behavioral Neuroscience*, *118*, 676–686.

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., & Gluck, M. A. (2005). The role of dopamine in cognitive sequence learning: evidence from Parkinson's disease. *Behavioral Brain Research*, *156*, 191–199.

Shohamy, D., Myers, C. E., Kalanithi, J., & Gluck, M. A. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neuroscience and Biobehavioural Reviews*, *32*, 219–236.

Sturges, J., & Whitfield, T.W.A. (1995). Locating basic colours in the Munsell Space. *Color research and application, 20*, 364-376.

Squire, L. R. (1992) Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195–231.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory, 82*, 171–177.

Squire L. R. (2009). Memory and brain systems: 1969–2009. *Journal of Neuroscience*, *29*, 12711–12716

Squire, L. R., & Zola, M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Science, U S A*, *93*, 13515–13522.

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, *40*, 385–98

Appendix

Order of 302 trials of the WPT and the corresponding actual outcome on each of them. In Experiments 1 and 2 only the first 200 trials were administered and the order is the typical one followed in previous studies (Prof. Meeter, personal communication), whereas in Experiment 3 all 302 trials were administered; the order for the additional 102 last trials was devised for the purposes of the present research. In italics are trials that are not included in the analysis, and in bold letters are trials that predict the less probable outcome.

Tria l	Pattern	Actual outcome	Trial	Pattern	Actual outcome
1	A	sun	30	C	sun
2	C	sun	31	A	sun
3	N	rain	**32**	**M**	**sun**
4	C	sun	33	A	sun
5	N	rain	34	N	rain
6	K	sun	35	G	sun
7	H	rain	36	M	rain
8	C	sun	37	G	sun
9	A	sun	38	C	sun
10	L	rain	39	L	rain
11	G	sun	40	H	rain
12	*I*	*rain*	41	B	sun
13	N	rain	42	N	rain
14	H	rain	43	L	rain
15	A	sun	**44**	**C**	**rain**
16	C	sun	45	H	rain
17	G	sun	46	L	rain
18	K	sun	47	H	rain
19	C	sun	48	N	rain
20	D	rain	49	A	sun
21	N	rain	50	H	rain
22	L	rain	51	D	rain
23	**N**	**sun**	52	C	sun
24	C	sun	**53**	**H**	**sun**
25	D	rain	54	L	rain
26	L	rain	55	N	rain
27	H	rain	56	B	sun
28	**K**	**rain**	57	N	rain
29	G	sun	**58**	**D**	**sun**

Trial	Pattern	Actual outcome	Trial	Pattern	Actual outcome
59	**B**	**rain**	**102**	**N**	**sun**
60	A	sun	103	J	rain
61	B	sun	**104**	**M**	**sun**
62	K	sun	*105*	*I*	*sun*
63	C	sun	106	N	rain
64	**K**	**rain**	107	G	sun
65	L	rain	108	E	sun
66	G	sun	*109*	*F*	*rain*
67	C	sun	110	L	rain
68	G	sun	**111**	**C**	**rain**
69	*I*	*rain*	*112*	*F*	*rain*
70	A	sun	113	C	sun
71	G	sun	114	N	rain
72	K	sun	115	E	sun
73	D	rain	116	A	sun
74	A	sun	117	D	rain
75	E	sun	118	C	sun
76	D	rain	119	E	sun
77	L	rain	120	K	sun
78	**G**	**rain**	121	H	rain
79	J	rain	122	A	sun
80	C	sun	123	G	sun
81	H	rain	**124**	**E**	**rain**
82	G	sun	**125**	**A**	**rain**
83	A	sun	126	D	rain
84	*F*	*sun*	127	N	rain
85	C	sun	**128**	**J**	**sun**
86	E	sun	129	C	sun
87	**B**	**rain**	130	L	rain
88	C	sun	*131*	*I*	*rain*
89	H	rain	132	G	sun
90	M	rain	133	C	sun
91	J	rain	134	G	sun
92	L	rain	135	H	rain
93	A	sun	136	A	sun
94	*I*	*sun*	137	C	sun
95	E	sun	138	G	sun
96	H	rain	139	C	sun
97	B	sun	**140**	**M**	**sun**
98	*F*	*sun*	*141*	*I*	*sun*
99	J	rain	**142**	**L**	**sun**
100	A	sun	143	B	sun

101	E	sun	144	L	rain
Tria l	Pattern	Actual outcome	Trial	Pattern	Actual outcome
145	H	rain	188	L	rain
146	L	rain	189	N	rain
147	H	rain	190	J	rain
148	G	sun	191	M	rain
149	J	rain	192	L	rain
150	B	sun	193	H	rain
151	L	rain	**194**	**M**	**sun**
152	**A**	**rain**	195	C	sun
153	M	rain	**196**	**H**	**sun**
154	**J**	**sun**	197	A	sun
155	E	sun	198	L	rain
156	G	sun	*199*	*F*	*sun*
157	C	sun	200	L	rain
158	H	rain	201	A	sun
159	L	rain	202	C	sun
160	N	rain	203	E	sun
161	**D**	**sun**	**204**	**K**	**rain**
162	**K**	**rain**	205	L	rain
163	E	sun	206	G	sun
164	N	rain	*207*	*F*	*sun*
165	B	sun	208	N	rain
166	C	sun	209	B	sun
167	J	rain	210	A	sun
168	**K**	**rain**	211	L	rain
169	E	sun	**212**	**E**	**rain**
170	**L**	**sun**	213	K	sun
171	N	rain	214	C	sun
172	L	rain	215	L	rain
173	N	rain	*216*	*I*	*sun*
174	L	rain	217	L	rain
175	J	rain	218	H	rain
176	**E**	**rain**	219	A	sun
177	L	rain	**220**	**D**	**sun**
178	G	sun	221	H	rain
179	J	rain	222	L	rain
180	**G**	**rain**	**223**	**N**	**sun**
181	L	rain	224	J	rain
182	H	rain	225	G	sun
183	*F*	*rain*	226	C	sun
184	C	sun	*227*	*F*	*rain*
185	M	rain	228	B	sun

186	J	rain	229	L	rain
187	A	sun	230	C	sun

Trial	Pattern	Actual outcome	Trial	Pattern	Actual outcome
231	N	Rain	264	D	rain
232	H	Rain	265	A	sun
233	A	Sun	*266*	*F*	*sun*
234	H	Rain	267	G	sun
235	**M**	**Sun**	268	N	rain
236	J	Rain	269	E	sun
237	**H**	**Sun**	270	J	rain
238	N	Rain	**271**	**K**	**rain**
239	G	Sun	272	N	rain
240	L	Rain	273	E	sun
241	H	Rain	281	A	sun
242	A	sun	282	L	rain
243	G	sun	**283**	**J**	**sun**
244	D	rain	284	G	sun
245	**L**	**sun**	285	H	rain
246	K	sun	286	C	sun
247	*F*	*rain*	287	M	rain
248	L	rain	**288**	**B**	**rain**
249	C	sun	289	E	sun
250	B	sun	290	C	sun
251	C	sun	291	J	rain
252	H	rain	292	M	rain
253	**A**	**rain**	293	C	sun
254	N	rain	*294*	*I*	*rain*
255	G	sun	295	D	rain
256	L	rain	296	C	sun
257	*I*	*rain*	297	A	sun
258	A	sun	298	N	rain
259	H	rain	299	E	sun
260	**G**	**rain**	**300**	**M**	**sun**
261	J	rain	301	L	rain
262	C	sun	**302**	**C**	**rain**
263	N	rain			