(2017), *Neuropsychology*, *31*, 862-876.

Quantitative Modeling of Category Learning Deficits in Various Patient Populations

J. Vincent Filoteo
Veterans Administration San Diego Healthcare System
Departments of Psychiatry and Neurosciences
University of California San Diego

W. Todd Maddox
Cognitive Design & Statistical Consulting, LLC

F. Gregory Ashby
Department of Psychological & Brain Sciences
University of California Santa Barbara

Objective: To provide a select review of our applications of quantitative modeling to highlight the utility of such approaches to better understand the neuropsychological deficits associated with various neurologic and psychiatric diseases.
Method: We review our work examining category learning in various patient populations, including individuals with basal ganglia disorders (Huntington's Disease and Parkinson's Disease), amnesia and Eating Disorders.
Results: Our review suggests that the use of quantitative models has enabled a better understand the learning deficits often observed in these conditions and has allowed us to form novel hypotheses about the neurobiological bases of their deficits.
Conclusions: We feel that the use of neurobiologically-inspired quantitative modeling holds great promise in neuropsychological assessment and that future clinical measures should incorporate the use of such models as part of their standard scoring. Appropriate studies need to be completed, however, to determine whether such modeling techniques adhere to the rigorous psychometric properties necessary for a valid and reliable application in a clinical setting.
Public Significance: This article reviews previous work highlighting the utility of quantitative modeling of cognitive processes in various patient populations. We feel that future clinical neuropsychological tests will incorporate such models in their standard scoring procedures.

This article selectively reviews our work using quantitative modeling to better understand learning deficits associated with various neurologic conditions. While it is understood that there have been many applications of modeling in various patient populations, the literature that we review allows us to describe a series of studies that have made use of a specific technique, the Perceptual Categorization Task (PCT; also called the General Recognition Randomization Technique; Ashby & Gott, 1988) that enables the consistent application of a set of quantitative models at the *individual* participant level, thus making these models highly relevant for evaluating individual differences in cognitive functioning. In this paper, we first provide a brief history of quantitative modeling in normal cognition and how past work has almost exclusively relied on analyzing aggregate data (data collapsed across all participants). We argue that this approach to modeling can lead to erroneous conclusions. We then briefly describe the PCT and how it has been used to test the Competition between Verbal and Implicit Systems (COVIS; Ashby, Alfonso-Reese, Turken, & Waldron, 1998) model of category learning. We next describe our applications of this quantitative modeling approach to the PCT with various patient populations and how these studies have helped support some of the most important aspects of the COVIS model. We then provide a brief description of some of the clinical applications of our quantitative modeling in various patient populations and how the application of this approach holds great promise for clinical applications. We conclude with a brief description of where we feel the field of neuropsychology might go in terms of future use of quantitative modeling in neuropsychological research and clinical work.

The Role of Quantitative Modeling in Neuropsychology

Statistical models, like ANOVA, model data without regard to the processes that produce that data. For example, a statistical model might assume that one group of scores is normally distributed and that two sets of scores are independent of each other. In contrast, process

models attempt to model the processes that produced the data. For example, signal detection models assume that hit and false alarm proportions are the result of a two-stage process (perception followed by a decision). Within neuropsychology, process models can be cognitive or neural (i.e., neurobiologically plausible). As their names suggest, cognitive process models describe the cognitive processes that mediate the behavior and neural process models describe the underlying neural processes based on what is known about the neurobiological systems being modeled.

In every scientific field the first models tend to be statistical, because statistical models require little or no knowledge of the processes producing the data. As knowledge about the phenomenon under study accumulates, process models appear. Neural models require more knowledge about the task than cognitive models, so not surprisingly, the first process models in neuropsychology were cognitive. The explosive growth during the past several decades in neuroscience, and especially cognitive neuroscience, has added enough knowledge that neural models are now also making important contributions to neuropsychology.

Statistical models can tell us that control and patient groups performed differently in a task, but they cannot tell us why. Process models have the more meaningful goal of explaining why the groups performed differently. For example, a statistical model might conclude that two groups performed differently in a certain task, whereas a cognitive model might conclude that the difference occurred because the groups perceived the stimuli in the same way, but the patient group made poorer decisions because of perseveration on a suboptimal strategy (e.g., Maddox & Filoteo, 2001). A neural model might add that the perseveration could possibly be due to reduced dopamine levels in the basal ganglia (e.g., Filoteo et al., 2014).

Another importance of process models is that they can provide accurate accounts of individual difference. In fact, process models go even further and warn that averaging data across participants, which is common in statistical modeling, often obscures the true structure of the data (Estes, 1956; Ashby, Maddox, & Lee, 1994; Maddox, 1999). Statistical models typically assume that variability in data is due to error (e.g., in measurement). If this is true, then averaging data across participants leads only to positive outcomes. Most importantly, error is reduced and because averaging brings the central limit theorem into play, the statistics become more normally distributed. However, process models warn that variability – especially variability across participants – can be due to process differences. Different participants might perform the same task using different cognitive strategies. The danger is then that their averaged data might not be characteristic of the cognitive processes used by any of the participants.

Estes (1956) cautioned about this problem many years ago. He showed that if every participant's accuracy jumps from 50% to 100% correct on one trial, but the trial on which this jump occurs varies across participants, then the resulting averaged learning curve will be gradually increasing. Thus, the averaged data suggest an incremental learning process, even though every participant displayed one-trial learning. Many other examples have been subsequently reported that document how averaging can change the psychological structure of data (Ashby, Maddox, & Lee, 1994; Maddox, 1999). For this reason, process models are typically applied separately to the data of each individual participant.

Category Learning as a Model Task in Neuropsychology

The same statistical model can often be applied to data from a wide variety of different tasks. In contrast, process models can only be applied to tasks that depend on the same processes. So although a good process model can tell us far more than any statistical model about why a particular patient group performed some task differently than the control group, the same process model will be applicable to a smaller domain of tasks than a statistical model like ANOVA. The most useful tasks for process modeling will share two properties. First, they must be relatively simple to ensure a reasonable understanding of the underlying cognitive and neural processes. Second, they should recruit a widespread set of neural structures. A task that depends on only a small set of brain regions might be easy to model, but it would only facilitate our understanding of patient groups with known dysfunction in one of the implicated regions.

One task that meets both criteria is perceptual category learning, which has been studied intensively for more than 50 years, and is known to depend on many different brain regions. Evidence of this latter fact can be seen in the many different neuropsychological populations who are impaired in some form of category learning. Included in this list are patients with Parkinson's disease (PD), schizophrenia, amnesia, Huntington's disease (HD), anorexia, and Alzheimer's disease. Furthermore, there are highly successful cognitive and neural process models of category learning.

The breakthrough that made accurate cognitive and neural process models of category learning possible were the theoretical proposals and then subsequent empirical evidence that humans have multiple learning systems that, for the most part, are functionally and anatomically distinct, evolved at different times for different purposes, and that learn in qualitatively different ways (Ashby et al., 1998; Ashby & Valentin, 2017a; Erickson & Kruschke,

1998; Nosofsky, Palmeri, & McKinley, 1994; Smith & Sloman, 1994). Roughly speaking, the category-learning systems shadow the brain's memory systems. Learning requires some sort of memory, and there is no clear reason why any memory system should be precluded from encoding memories about categories, so it is likely that there are as many category-learning systems as there are memory systems (Ashby & O'Brien, 2005). Thus, every patient population with any kind of memory deficit is likely to be impaired in some form of category learning.

Category-learning studies with neuropsychological populations have primarily focused on two different types of category learning: implicit and explicit (see Figure 1). While this is somewhat of a simplistic distinction, and it is obvious that overlapping neural systems contribute to both implicit and explicit learning (e.g., Poldrack & Packard, 2003), past behavioral and functional neuroimaging work with normal participants and various patient populations provides extensive evidence that implicit and explicit category learning are mediated differently (Ashby *et al.*, 1998; Filoteo, Maddox, Simmons *et al.*, 2005; Filoteo, Simmons, Zeithamova, Maddox, & Paulus, 2006; Knowlton & Squire, 1993; Maddox, Filoteo, Hejl, & Ing, 2004; Maddox, Filoteo, & Lauritzen, 2007; Maddox, Filoteo, Lauritzen, Connally, & Hejl, 2005; Nomura *et al.*, 2007; E. E. Smith, Patalano, & Jonides, 1998; E. E. Smith & Sloman, 1994). Explicit category learning is dependent on hypothesis generation, logical reasoning, working memory and executive attention. Tasks that measure explicit category learning are often referred to as *rule-based* (RB) tasks, because there is typically a verbalizable "rule" that defines category membership. The Wisconsin Card Sorting Task is perhaps the most widely known example (Heaton, 1981). Evidence suggests that rule-based category learning is mediated within an anterior brain network that includes the dorsolateral frontal lobes and the anterior caudate nucleus (Ashby *et al.*, 1998).

FIGURE 1 HERE

In contrast, there are several forms of implicit category learning in which a participant can learn categories without having any conscious awareness of a categorization strategy. Some of the tasks that have been used to examine implicit category learning in patient populations are *information-integration* (II) tasks, *prototype-distortion* tasks, and *artificial-grammar* tasks. II category-learning tasks, which require a similarity-based decision strategy that is impossible to describe verbally, are thought to be learned via a procedural system that uses a dopamine-mediated reward signal in the basal ganglia to associate a perceptual stimulus with a motor response (Ashby & Waldron, 1999). The key structures are the posterior caudate/putamen and premotor cortex (e.g., supplementary motor area) (Nomura *et al.*, 2007; Waldschmidt & Ashby, 2011). As their names imply, prototype-distortion tasks require participants to identify distortions of a prototype pattern and artificial grammar learning tasks require participants to decide whether a given letter string is consistent with some novel, unknown grammar. Both tasks are thought to be learned primarily through a perceptual priming system that is mediated within posterior visual cortices (Casale & Ashby, 2008; Reber & Squire, 1999; Reber, Stark, & Squire, 1998), although other neural systems are also involved (Forkstam, Hagoort, Fernandez, Ingvar, & Petersson et al., 2006; Skosnik et al., 2002).

The majority of our category-learning studies used the PCT first developed by Ashby and Gott (1988), in which individuals are presented with simple stimuli and asked to learn to categorize them into distinct groups. The stimuli often consist of lines that vary in length and orientation or Gabor patches that vary in orientation and spatial frequency (see Figures 2 and 3). In this task, participants are presented with a stimulus and are asked to categorize it into Category A or Category B. Once a response is made, the participant is given immediate corrective feedback. Each category is defined as a probability distribution – often a multivariate normal distribution – and prior to the start of the experiment the category exemplars are created by drawing a large number of random samples from each of these distributions. For a complete description of all the methods needed to run an experiment of this type, see Ashby and Valentin (in press-b).

Figure 2 shows samples from two category distributions in which the stimuli are Gabor patches that vary in orientation and spatial frequency. Each stimulus is represented as a point in two-dimensional space. In Figure 2, the x-axis represents the spatial frequencies (i.e., widths) of the dark and light bars and the y-axis represents their orientation. Black squares represent exemplars of Category A and open circles represent exemplars of Category B. The arrows in Figure 2 link a sample stimulus to its representation in this two-dimensional stimulus space. In these studies, a single categorization strategy maximizes accuracy. The nature of this strategy is determined by the relationship between the two category distributions, and thus, by the two stimulus attributes. When the categories are each defined by a multivariate normal distribution (as in Figures 2 and 3), the optimal strategy is always equivalent to using a linear or quadratic decision boundary that divides the stimulus space into separate category A and B response regions. The solid line in Figure 2 denotes this optimal decision bound. A participant who responds A to any stimulus falling to the left of this bound and B to any stimulus falling to the right will maximize long-run accuracy.

FIGURE 2 HERE

A major advantage in using the PCT is that it allows us to construct tasks that require participants to learn different types of categorization strategies, either implicit or explicit, by simply changing the distribution of the stimuli within the categories. Specifically, the task depicted in Figure 2 requires an explicit rule because the optimal strategy (depicted as the solid line) is easily verbalized. In essence, optimal performance requires that the participant learn to attend to only the spatial frequency of the stimuli and identify the value of bar width that best separates the two categories. This rule is easy to describe verbally: "assign stimuli with wide bars to Category A, and stimuli with narrow bars to Category B". For this reason, Figure 2 depicts a rule-based task. In contrast, Figure 3 depicts two examples of II tasks that require a procedural categorization strategy – one in which the optimal decision bound is linear (Figure 3A) and one in which it is nonlinear (i.e., quadratic; Figure 3B). In this example, the optimal strategy that defines category membership is based on a relationship between the length and the orientation of the line stimuli (that is, information from the two dimensions must be integrated). Because these stimuli are in non-commensurable physical units (length and orientation), it is difficult to verbalize an optimal rule of this nature, and thus learning has to occur at an implicit level. In these examples, the rule depicted in Figure 3A is based on a linear combination of the two stimulus dimensions, whereas the rule depicted in Figure 3B is based on a nonlinear combination of the two dimensions. A large body of research has demonstrated that learning in RB and II tasks is dissociable (Ashby & Maddox, 2011).

FIGURE 3 HERE

COVIS Model of Category Learning

The past 20 years has seen enormous advances in our understanding of the neural processes that mediate learning in II and RB tasks. The most comprehensive and best tested neural theory is the COVIS theory of category learning (Ashby et al., 1998; Ashby & Crossley, 2011; Ashby, Ennis, & Spiering, 2007; Ashby & Waldron, 1999; Cantwell, Crossley, & Ashby, 2015). Briefly, COVIS postulates two systems that compete throughout learning – a frontal-based system that learns explicit rules and depends on declarative memory systems and a basal ganglia-mediated procedural-learning system. The procedural system is phylogenetically older. It can learn a wide variety of category structures, but it learns in a slow incremental fashion and is highly dependent on reliable and immediate feedback. In contrast, the explicit rule-learning system can learn a fairly small set of category structures quickly – specifically, those structures in which the contrasting categories can be separated by simple explicit rules. So COVIS assumes that the explicit, rule-learning system dominates in RB tasks. But COVIS also assumes that the explicit system is unable to learn the optimal strategy in II tasks, and so must pass control to the procedural-learning system.

The key structures in the COVIS explicit, rule-learning model include the anterior cingulate, the prefrontal cortex, the head of the caudate nucleus, and the hippocampus. The model postulates that rule learning includes a number of separate sub-processes, including: generating or selecting new candidate rules, maintaining candidate rules in working memory, applying these rules, using feedback to update rule salience, and switching attention away from a discredited rule to a newly selected rule. COVIS predicts that rule maintenance, rule selection, and rule switching are all facilitated by dopamine.

The key structure in the COVIS procedural-learning system is the striatum. COVIS assumes that arbitrary stimulus-response associations of the type required for success in II tasks are learned via incremental changes in synaptic plasticity at cortical-striatal synapses that are facilitated by a form of dopamine-mediated reinforcement learning. The terminal site of learning is presumed to lie somewhere in premotor cortex, which explains for example, why switching the location of the response keys interferes with performance in II tasks, but not in RB tasks (Ashby, Ell, & Waldron, 2003; Maddox, Bohil, & Ing, 2004).

A complete review of COVIS is beyond the scope of this article. For a recent detailed review of the neuroscience of COVIS, see Ashby and Valentin (2017-a), and for a detailed computational description, see Ashby, Paul, and Maddox (2011). In COVIS, dopamine has different effects on the rule-learning and procedural systems. In the rule-learning system, COVIS assumes that rule selection and rule switching are both impaired if brain dopamine levels decrease, but the model assumes that selection depends on cortical dopamine levels, whereas switching depends on basal ganglia dopamine levels (Ashby et al., 1998). In the procedural system, dopamine plays a crucial role in learning: it provides the reward signal required for reinforcement learning. A reduction in tonic levels of dopamine or in the dynamic range of phasic dopamine bursts impairs the ability of the procedural system to learn stimulus-response associations.

The gradual loss of dopamine producing cells in Parkinson's disease (PD) leads to dysfunction in frontal and basal ganglia systems. Thus, PD was identified early on as an ideal patient population to examine the neurological plausibility of COVIS, given the role that COVIS proposes for frontal brain regions, basal ganglia structures, and dopamine in the learning of RB and II categories.

Cognitive Process Modeling of Neuropsychological Deficits

Another major advantage of examining performance of PD and other patient populations on the PCT, and the one most relevant to this article, is that it readily lends itself to the application of sophisticated process models. The earliest applications used cognitive models (e.g., Filoteo, & Maddox, 1999; Filoteo, Maddox, Salmon, & Song, 2005; Maddox, & Filoteo, 2001; Maddox, Filoteo, Delis, & Salmon, 1996), but more recently neural process models have also been applied successfully to several different patient groups (Ashby et al., 1998; Filoteo et al., 2014; Hélie, Paul, & Ashby, 2012a, 2012b). This section describes our work using cognitive process models to better understand deficits associated with a variety of different neuropsychological conditions.

Our cognitive process modeling has almost exclusively used decision bound models (Ashby, 1992; Maddox & Ashby, 1993). Decision bound models assume that participants partition the perceptual space into response regions. On every trial, the participant determines which region the percept is in, and then emits the associated response. Two different types of decision-bound models are typically fit to the responses of each individual participant: models that assume an explicit rule-learning strategy and models that assume an implicit procedural strategy. It is also common to fit other models that assume the participant guesses at random on every trial. Decision bound models have two types of parameters – those that describe the shape and location of the decision bound, and at least one variance parameter that describes trial-by-trial variability in perceptual and decisional processing. In addition, two statistics are used to assess the quality of fit (Ashby, 1992). First, the overall *goodness-of-fit* (or fit) value of the model measures how well the model accounted for the participant's responses (with a low value signaling a good fit). Second, the numerical value of the variance estimate, hereafter referred to as the *application variability* value, measures how consistently the participant applied the same strategy on a trial-by-trial basis (with a low value signaling a consistent strategy).

Decision bound models are derived from general recognition theory (Ashby & Townsend, 1986), which is a multivariate generalization of signal detection theory (e.g., Macmillan & Creelman, 2004; Wickens, 2002). As in signal detection theory, decision bound models separate the categorization response into separate perceptual and decisional components. Therefore, a comparison of the parameter estimates from the best-fitting models can reveal whether the deficits associated with a specific neuropsychological disorder primarily affect how the patients perceive the world or how they learn and make decisions. In addition, comparing fits of explicit rule-learning models, implicit procedural-learning models, and guessing models allow one to identify the general strategy a participant used to perform a given task (e.g., implicit vs. explicit). This is necessary because it is sometimes the case that a participant will attempt to use one approach to solve a task, such as an explicit approach, despite the fact that another approach, such as an implicit approach, is more optimal and would lead to higher accuracy. For a more detailed description of decision-bound models, and a step-by-step description of how to apply these models to categorization data, see Ashby and Valentin (in press-b).

Modeling of Implicit Category Learning

In two separate studies (Maddox & Filoteo, 2001; Filoteo, Maddox, Salmon, et al., 2005) we compared category learning in a group of PD patients to a group of age, education, and gender-matched healthy control (HC) participants using II tasks in which the optimal decision bounds were linear and nonlinear, and similar to the bounds depicted in Figures 3a and 3b, respectively. Accuracy results for Maddox and Filoteo (2001), which are depicted in Figure 4, demonstrated that the PD patients performed as well as HC participants in learning the linearly separable categories, but they were impaired in learning the categories that were nonlinearly separable.

FIGURE 4 HERE

To examine PD patients' deficit in the nonlinear II condition in the Maddox and Filoteo (2001), we next applied a decision-bound model (called the general quadratic categorization model) that assumed the participant used a categorization strategy that was based on a suboptimal decision bound that was nevertheless of the same form as the optimal bound (i.e., so quadratic, but with a different shape from the optimal bound depicted in Figure 3b; Ashby, 1992; Maddox &Ashby, 1993). For each participant, we computed the goodness-of-fit of the model and the application variability value. The average of these indices are displayed in Figure 5. Interestingly, the difference in fit values for the PD patients and HC participants remained discrepant across the 600 trials, whereas the application variability values were less discrepant across the 600 trials. These results suggest that PD patients are impaired at learning the procedural strategy, as indicated by the consistent difference between the PD patients and HC participants in fit values across the trials of the experiment, but once they learned a strategy, the PD patients were able to consistently apply that strategy across the trials of the task.

FIGURE 5 HERE

In a subsequent study with PD patients (Filoteo, Maddox, Salmon, et al., 2005) we again administered linear and nonlinear II tasks to PD patients and HC participants using the category distributions depicted in

Figure 3. The accuracy results from this study again suggested that PD patients are impaired in learning nonlinearly separable II categories, but not in learning II categories that are linearly separable (see Figure 6). This latter finding is consistent with Maddox and Filoteo (2001) as well as another previous study of ours (Ashby, Noble, Filoteo, Waldron, & Ell, 2003) that demonstrated PD patients were not impaired in learning II categories that were defined by a linear relationship between stimulus dimensions. Taken together, these results suggest that the basal ganglia are involved in learning complex categorization strategies and that the deficit in the nonlinear II task could be due to PD patients having greater difficulty in learning to categorize perceptually dissimilar stimuli into the same category. This latter possibility was also supported by a recent study we conducted (Filoteo & Maddox, 2014).

FIGURE 6 HERE

In regard to the modeling of the data from Filoteo, Maddox, Salmon et al. (2005), we examined the fit values for the models that best fit each participant's data in the final block of trials. The only difference between the two groups were the fit values from the nonlinear II task, with the PD patients fit values being significantly greater than the HC participants, a finding that indicates that the strategy used by the PD patients was less optimal than the HC participants, but they applied that process just as consistently as HC participants. These findings replicate those from our initial II study with PD patients (Maddox & Filoteo, 2001).

We also applied several decision-bound models to the data from Filoteo, Maddox, Salmon et al. (2005) to determine if participants learned the categories using either an implicit, procedural strategy or an explicit, rule strategy. Each model was fit to the final block of trials, which tend to be the most stable responses. The results indicated no statistically significant differences between the percentages of PD patients and HC participants who used a procedural approach in the linear condition (55% vs. 65%, respectively) or the nonlinear condition (80% for both groups). We next compared the accuracy rates for the final block of trials in the linear and nonlinear conditions for only those PD patients and HC participants who used a procedural approach. The results are depicted in Figure 7 and show that for only those PD patients and HC participants who used a procedural approach, the groups did not differ in the linear II condition, but PD patients were significantly less accurate than the HC participants in the nonlinear II condition. These results indicate that PD patients who adopted a procedural approach in the nonlinear condition were impaired relative to those HC participants who also adopted a procedural approach, and they suggest that an impairment in II category learning significantly contributed to PD patients' deficits in the overall accuracy analysis. These findings are important in that they demonstrate that PD patients are actually impaired in learning nonlinear II tasks even when they used a procedural approach to learning the task. That is, their deficit was not due to using the wrong approach, but rather to using the correct approach less accurately.

FIGURE 7

It is also important to note that patients with amnesia are not impaired in learning nonlinear II tasks (Filoteo, Maddox, & Davis, 2001a). In this study, we examined two amnesiac patients, one who developed amnesia following bilateral medial temporal lobe damage secondary to encephalitis (patient JW), and a second patient who developed amnesia following white matter damage secondary to a drug-binge that resulted in a vascular event (patient PK). Accuracy rates and decision-bound models indicated that the two amnesiacs were normal in learning the nonlinear II task relative to HC participants. Furthermore, patient JW was able to retain the rule over a 24-hour period. These results indicate that memory systems involved in explicit learning are not intricately involved in learning nonlinear II tasks.

Another form of category learning that has been studied in patients with PD is probabilistic learning, in which a set of stimuli is probabilistically related to one of two outcomes. The most popular task that has been used with PD patients is the Weather Prediction Task (WPT), which requires subjects to learn to categorize stimuli (consisting of various cue combinations) that are probabilistically associated with one of two categorical outcomes – 'rain' or 'sunshine' (Gluck, Oliver, & Myers, 1996). Most studies have demonstrated that PD patients show some form of impairment on the WPT relative to HC participants (Knowlton, Mangels, & Squire, 1996; Shohamy, Myers, Grossman *et al.*, 2004; Witt, Nuhsman, & Deuschl, 2002; but see Moody, Bookheimer, Vanek, & Knowlton, 2004; Price, 2005).

To examine the nature of PD patients' deficits on the WPT, Gluck and colleagues (Gluck, Shohamy, & Myers, 2002) quantitatively instantiated several different strategic approaches one could use when performing the WPT and applied this strategy analysis to PD patients' performances on this task (Shohamy, Myers, Grossman *et al.*, 2004; Shohamy, Myers, Onlaor, & Gluck., 2004). The results indicated that both PD patients and HC participants tended to learn the WPT early on by memorizing stimuli with only a single cue present (referred to as a singleton strategy). As learning progressed, however, the majority of HC participants tended to switch to 'multi-cue' strategies that required the integration of multiple cues within the display. In contrast, the PD patients tended to continue to use a singleton strategy that they had adopted during the early

part of learning and failed to switch away to the more advantageous multi-cue approach, which is consistent with other studies that demonstrated an association between WPT performance and the number of perseverative errors on the WCST (Knowlton *et al.*, 1996; Price, 2005; but see Shohamy, Myers, Grossman et al., 2004 for an alternative explanation).

Furthermore, Shohamy, Myers, Onlaor *et al.* (2004) found that PD patients and HC who switched to a multi-cue strategy did not differ on the WPT in terms of accuracy, suggesting that when patients can change to a more efficient strategy on the WPT, they are able to apply it just as accurately as controls. This finding is in contrast to Filoteo, Maddox, Salmon et al. (2005) who showed that PD patients who used a procedural approach were still impaired in terms of accuracy in learning a nonlinear II task, as compared to HC participants who also used a procedural approach. Taken together, the approaches applied by our group, and in the work of Shohamy and Gluck, demonstrate the utility of quantitative modeling for a more fine-grained analysis of the cognitive processes involved in category learning.

Modeling of Explicit Category Learning

In several previous studies, we examined PD patient's ability to learn RB tasks where the optimal rule was similar to that shown in Figure 2 (Ashby, Noble, *et al.*, 2003; Filoteo, Maddox, Ing, Zizak, & Song, 2007; Maddox & Filoteo, 2001; Maddox, Aparicio, Merchant, & Ivry, 2005). As noted above, optimal responding required the participant to categorize the stimulus into one category if the widths of the bars from the Gabor patch were narrow or into the other category if the widths of the bars were wide. In this case, the orientation of the bars did not determine category membership, although the stimuli varied on this dimension on each trial. In each of the studies that examined this type of RB task, we found that PD patients were impaired, but only when there was variation on the stimulus dimension that was not relevant for categorizing the stimuli. These results indicate that the ability of PD patients to learn the explicit categories was impacted to a much greater extent than HC participants as the number of varying irrelevant dimensions increased (c.f., Filoteo, Maddox, Ing, Zizak, & Song, 2005), suggesting that deficits in selective attention might contribute to the PD patients' impairment in explicit category learning.

In addition to deficits in selective attention, PD patients have also been shown to be impaired in other processes known to be required for learning RB tasks. For example, in learning explicit rules, participants must generate hypotheses regarding the possible rule, test such hypotheses using feedback, switch to a new hypothesis if the one currently in use is not correct, and keep track of those hypotheses that either did not work or are currently working. These processes rely to a large extent on working memory, and given that PD patients have been shown to be impaired in this process (Gilbert, Belleville, Bherer, & Chouinard, 2005; Owen *et al.*, 1993; Owen, Iddon, Hodges, Summers, & Robbins, 1997; Postle, Jonides, Smith, Corkin, & Growden, 1997), it is possible that deficits in working memory might also contribute to PD patients' impairment in learning explicit categorization rules. To examine this issue, we conducted a study (Filoteo, Maddox, Ing, et al., 2007) in which PD patients and HC participants were asked to learn three explicit category structures: the rule depicted in Figure 2, which as noted previously is an RB task that emphasizes selective attention, and the other two rules that are depicted in Figure 8, which emphasize working memory. In the *conjunctive* RB condition depicted in Figure 8A, optimal responding required the subject to respond "A" if the stimulus was more vertical *and* had narrow bars, and otherwise to respond "B". Note that this strategy combines independent decisions about the two features and is highly verbalizable. As such, this task is considered to be explicit. The optimal rule is depicted by the solid horizontal and vertical lines in Figure 8A. In the *disjunctive* condition depicted in Figure 8B, the optimal rule required that the subject respond "A" if the stimulus was more vertical and had narrow bars *or* if the stimulus was more horizontal and had wide bars, and to respond "B" if the stimulus was more vertical and had wide bars *or* if the stimulus was more horizontal and had narrow bars. The optimal rule is depicted by the solid horizontal and vertical lines in Figure 8B. Although optimal responding in both the conjunctive and disjunctive tasks required participants to use a verbalizable combination of the two stimulus dimensions, the two tasks likely emphasize working memory to a different degree. Specifically, the logical expression associated with the disjunctive rule is much longer than the logical expression associated with the conjunctive rule, and therefore should require greater working memory. Thus, a comparison of PD patients' performances in the conjunctive and disjunctive conditions could help determine whether working memory deficits might also contribute to PD patients' explicit category learning deficits.

FIGURE 8 HERE

Figure 9 displays the accuracy results from the three conditions. As can be seen, PD patients demonstrated a large impairment on the one-dimensional explicit condition (Figure 9a), replicating previous findings (Ashby *et al.*, 2003; Maddox, Aparicio *et al.*, 2005). In contrast, the patients were not impaired in the conjunctive condition (Figure 9b) or the disjunctive condition (Figure 9c). Importantly, both groups displayed less learning in the disjunctive condition than the conjunctive condition,

which was likely due to the greater working memory requirements of the former task. The pattern of PD patients' performance suggests that the explicit deficit exhibited by these patients in past studies is likely related to impairment in selective attention, but not working memory.

FIGURE 9

We next applied several procedural-learning and rule-learning models to the data, and as might be expected, we found that in the conjunctive and disjunctive conditions there were no group differences in percentage of PD patients or HC participants who used a rule-learning process in the last block of trials (conjunctive condition: PD = 80%, HC = 73%; disjunctive condition: both groups = 100%) and there were no differences between the two groups in their fit values or variability estimates that best accounted for their data. In contrast, in the single-dimension condition, although both groups used a rule-learning process (PD = 92%, HC = 75%), the fit values indicated that PD patients were less able to use the optimal rule when learning to categorize the stimuli compared to the HC participants, but the two groups did not differ on the variability estimates.

Clinical Applications of Quantitative Modeling of Category Learning in PD

Our previous work highlights that PD patients can be impaired on both II and RB tasks relative to HC participants, but for very different reasons. In the case of RB tasks, it appears that selective attention deficits play a major role in the impairments we observed in PD, whereas in the case of II tasks it appears that PD patients are impaired in learning more complex procedural strategies (i.e., when the optimal approach is nonlinear). Although these results are interesting and tell us to some extent under what circumstances the basal ganglia play a role in category learning, they tell us very little about the potential clinical utility of examining category learning in PD. An important step in doing so is to determine both the cross sectional sensitivity and specificity of category learning deficits in PD, as well as the predictability of future cognitive decline based on current category learning performances.

To further determine the potential clinical utility of PD patients' category learning deficits, we re-examined their accuracy performances on the nonlinear II task described above using several different approaches and found the following in our PD samples: (1) nonlinear II accuracy was more sensitive than traditional measures of neuropsychological functioning (e.g., WCST and verbal fluency) in differentiating between PD patients and HC participants cross-sectionally, (2) accuracy on the nonlinear II task had a .91 positive predictive value (i.e., the probability that an individual has PD given they are impaired on the task), and a .74 negative predictive value (i.e., the probability that an individual does not have PD given they were not impaired on the task) in making this distinction, and (3) PD patients' accuracy on the nonlinear II task was highly predictive (r = -.78; 61% of the variance accounted for) of future decline on a measure of global cognitive functioning (total score on the Mattis Dementia Rating Scale; Mattis, 1988), even after age, gender, motor impairment, mood, baseline performance on the MDRS, and performance on the WCST were taken into consideration (Filoteo, Maddox, Song, & Salmon, 2007).

We also examined whether our quantitative modeling analyses would provide any additional predictive information regarding global cognitive decline. We found that PD patients whose responses on the nonlinear II task were best fit by one of the procedural-learning models declined less on the MDRS than those whose data were best fit by a rule-learning model. This difference is depicted in Figure 10. Most importantly, we determined whether the inclusion of the decision-bound models could help predict decline on the MDRS above and beyond what was predicted by accuracy performance alone. As noted above, final-block accuracy in the nonlinear II condition predicted 61% of the variance associated with future decline on the MDRS. To examine this issue, we conducted a stepwise regression analysis in which we predicted change on the MDRS by first entering final-block accuracy and then in the next step entering whether a patient's performance was best fit by an implicit or an explicit model. The inclusion of this latter variable predicted a significant additional 15% of the variance above and beyond the 61% predicted by accuracy level alone. Thus, using a single II category-learning task, we were able to predict 76% of the total variance associated with future cognitive decline in a nondemented PD sample after a relatively brief period of time (just 1.6 years). These results clearly establish the clinical utility for the use of quantitative modeling for a better prediction of global cognitive decline in nondemented PD patients.

FIGURE 10 HERE

Another important question in determining the clinical utility of any assessment approach is to determine whether that approach has any specificity in differentiating among various patient populations. To examine this issue, we developed a modified version of COVIS (Maddox, Filoteo, & Zeithamova, 2010) and applied it to PD patients' nonlinear II data as well as a data set obtained from patients with Huntington's disease (HD; Filoteo, Maddox, & Davis, 2001b), a patient population that is well known to have basal ganglia pathology. In this application, the models suggest that the locus of HD patients' nonlinear II deficit is in their increased reliance on rule-learning strategies. Note that this finding is in contrast to that for PD patients whose

deficit was in the application of sub-optimal procedural-based strategies. This finding is important because it allows us to pinpoint differences in the PD and HD patients' deficits in performing an II task, which could not have been done based on accuracy alone.

A final potentially important application of quantitative modeling is to identify specific sub-processes that might underlie any patient population's impairment on a given task. Recently, we used a neural process model to examine cognitive deficits in individuals with Anorexia Nervosa (AN) (Filoteo et al., 2014). AN is also thought to impact the basal ganglia and disrupt several neurocognitive processes that we have demonstrated to be impaired in patients with PD, such as cognitive set shifting (Roberts, Tchanturia, Stahl, Southgate, & Treasure, 2007; Roberts, Tchanturia, & Treasure, 2010; Shott et al., 2012; Steinglass, Walsh, & Stern, 2006; Tchanturia, Morris, Surguladze, & Treasure, 2002, Tchanturia et al., 2011). As we demonstrated in our past work, the deficit on such tasks in PD appear to be related to deficits in selective attention and not set-shifting, per se, but the mechanisms underlying set-shifting deficits in AN had not been studied in any detail.

To address this issue in AN, we administered an RB category-learning task in which participants were shown multi-dimensional stimuli, such as a cartoon version of a house, and asked to categorize each image into one of two categories. The RB rule that determined category membership was based on a single dimension and was highly verbalizable (e.g., if the roof of the house is flat it belongs to Category 1, but if the shape of the roof is a triangle, then it belongs to Category 2). Participants were given immediate feedback as to the accuracy of each response. After 80 trials, the rule shifted to a different dimension without the participant being told (e.g., now the relevant rule was if there was a tree in front of the house, the stimulus belonged to Category 1, but if there was not a tree in front of the house the stimulus belonged to Category 2). The accuracy results are shown in Figure 11 and demonstrated that, relative to their controls, the AN participants demonstrated greater learning prior to the rule shift, but when the rule shifted, the controls were better able to learn the new rule, whereas the AN participants had greater difficulty. These accuracy findings suggested hyper-learning prior to the rule shift and impaired set-shifting after the rule shift in the AN group.

FIGURE 11 HERE

The application of neural modeling to AN is philosophically different than to PD. In the case of PD, we have a clear understanding of the underlying neural etiology – namely, the death of dopamine-producing neurons. So the primary reason to fit a neural model to data from a task where the cognitive deficits of PD patients are well understood is to test the model. If the model is valid, then reducing dopamine levels in the model should produce cognitive deficits similar to those seen in PD patients, without the need to make any other modifications to the model. Once validated, the model can then be used for other more interesting applications, such as predicting how PD patients should perform in some novel task, or investigating the efficacy of some novel PD treatment.

On the other hand, the etiology of AN is less clear. In this case, the model can be used to better understand the neural and cognitive differences between individuals with AN and HC participants. The idea is to fit the model to data from both groups and examine which parameters need to be manipulated to account for the group differences. When Filoteo et al. (2014) followed this strategy, they found that the model gave good fits to all aspects of both data sets under the following conditions. First, the AN group showed hypersensitivity to negative feedback, which accounted for the hyper-learning on the first category structure (but did not cause the model to predict hyper-learning on subsequent structures). Second, the AN group had reduced brain dopamine levels, which caused set-shifting deficits (i.e., perseveration). Both of these conditions have been previously suspected in AN (Bailer et al., 2013; Frank et al., 2005; Harrison, O'Brien, Lopez & Treasure, 2010; Harrison, Treasure, & Smillie, 2011; Jappe et al., 2011), so in this case the neural modeling provided a more rigorous test of existing theories.

This latter application of a neural model to the behavioral data of AN patients highlights the utility of this approach. By identifying what parameters need to be altered to replicate behavioral findings in AN, we can perhaps more precisely pinpoint the underlying neurocognitive mechanisms that are impaired. Furthermore, if such mechanisms are found to be play a role in the pathogenesis of the disease, the model can then be used to determine if different interventions have any behavioral effects on those model parameters thought to represent those processes. For example, if a behavioral intervention targeted at decreasing AN patients' sensitivity to punishment can be identified, then we would expect a decrease in their hyper-learning prior to a rule shift, whereas a trial of dopaminergic medication might improve their ability to shift set after the rule change.

Future Directions

We conclude this article with a few thoughts on the possible future role of quantitative modeling in neuropsychology and psychiatry. First, we feel that the application of modeling in patient populations will continue to provide a greater understanding of the specific cognitive and neural processes that lead to impaired

performance on a given cognitive task. For example, our application of a neural processing model to the hyper-learning and impaired set-shifting in AN allowed us to formulate and test more detailed hypotheses regarding the underlying causes of these phenomena. Second, we demonstrated that quantitative modeling can be applied successfully at the individual participant level, which is an obvious prerequisite for their use in clinical assessment. As an example, signal-detection models have been incorporated into clinical neuropsychological tests of recognition memory (e.g., the California Verbal Learning Test-2; Delis et al., 2000). We feel that this is a future direction for process models such as those described in this article, and that the data provided by these models could potentially be part of the standard scoring of clinical neuropsychological tests. So, in addition to accuracy or response time, clinical neuropsychologists of the future should be able to use process models to gain more insight about the specific processes impaired in their patients. As such, a major step forward will be for clinical studies to determine the psychometric properties of the data provided by process models to make certain that they meet the rigorous requirements needed to assess neuropsychological functioning in clinical populations.

Acknowledgements

This research was supported, in party, by a VA Merit Award 5I01CX000813-04 to JVF and National Institute of Health Grant 2R01MH0637 to FGA. Correspondence concerning this article should be addressed to J. Vincent Filoteo, UCSD, VASDHS 116-B, 3350 La Jolla Village Dr., San Diego, CA 92161 (email: vfiloteo@ucsd.edu).

References

Ashby, F. G. (1992). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449-483). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105*(3), 442-481.

Ashby, F. G., & Crossley, M. J. (2011). A computational model of how cholinergic interneurons protect striatal-dependent learning. *Journal of Cognitive Neuroscience, 23*, 1549-1566.

Ashby, F. G., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition, 31*, 1114-1125.

Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychological Review, 114*, 632-656.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning Memory & Cognition, 14*(1), 33-53.

Ashby, F.G. & Maddox, W.T. (2011). Human category learning. *Annals of the New York Academy of Sciences, 1224*, 147-161.

Ashby, F. G., Maddox, W. T., & Lee, W. W. (1994). On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science, 5*, 144–151.

Ashby, F. G., Noble, S., Filoteo, J. V., Waldron, E. M., & Ell, S. W. (2003). Category learning deficits in Parkinson's disease. *Neuropsychology, 17*(1), 115-124.

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Science, 2*, 83-89.

Ashby, F. G., Paul, E. J., & Maddox, W. T. (2011). COVIS. In E. M. Pothos & A.J. Wills (Eds.), *Formal approaches in categorization* (pp. 65-87). New York: Cambridge University Press.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93*, 154-179.

Ashby, F. G., & Valentin, V. V. (2017- a). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen and C. Lefebvre (Eds.), *Categorization in cognitive science, 2nd Edition,* (pp. 157-188). New York: Elsevier.

Ashby, F. G., & Valentin, V. V. (in press-b). The categorization experiment: Experimental design and data analysis. In E. J. Wagenmakers & J. T. Wixted (Eds.), *Stevens' handbook of experimental psychology and cognitive neuroscience, Fourth Edition, Volume Five: Methodology*. New York: Wiley.

Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin & Review, 6*(3), 363-378.

Bailer UF, Frank GK, Price JC, Meltzer CC, Becker C, Mathis CA, et al. Interaction between serotonin transporter and dopamine D2/D3 receptor radioligand measures is associated with harm avoidant symptoms in anorexia and bulimia nervosa. Psychiatry Research. 2013; 211:160–1.

Cantwell, G., Crossley, M. J., & Ashby, F. G. (2015). Multiple stages of learning in perceptual categorization: evidence and neurocomputational theory. *Psychonomic Bulletin & Review, 22*, 1598–1613.

Casale, M. B., & Ashby, F. G. (2008). A role for the perceptual representation memory system in category learning. *Perception & Psychophysics, 70*, 983-999.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (2000). California Verbal Learning Test–Second Edition, Adult Version manual. San Antonio, TX: The Psychological Corporation.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107-140.

Estes, W. K. (1956). The problem of inference from curves based on group data. P*sychological Bulletin*, *53*, 134–140.

Filoteo, J. V., & Maddox, W. T. (1999). Quantitative modeling of visual attention processes in patients with Parkinson's disease: effects of stimulus integrality on selective attention and dimensional integration. *Neuropsychology, 13*(2), 206-222.

Filoteo, J. V., & Maddox, W. T. (2014). Procedural-based category learning in patients with Parkinson's disease: Impact of category number and category continuity. *Frontiers in Systems Neuroscience, 8.*

Filoteo, J.V., Maddox, W.T., & Davis, J.D. (2001a). Quantitative modeling of category learning in amnesic patients. *Journal of the International Neuropsychological Society, 7*, 1–19.

Filoteo, J. V., Maddox, W. T., & Davis, J. D. (2001b). A possible role of the striatum in linear and nonlinear category learning: Evidence from patients with Huntington's disease. *Behavioral Neuroscience*, 115(4), 786-798.

Filoteo, J. V., Maddox, W. T., Ing, A. D., & Song, D. D. (2007). Characterizing rule-based category learning deficits in patients with Parkinson's disease. *Neuropsychologia, 45*(2), 305-320.

Filoteo, J. V., Maddox, W. T., Ing, A. D., Zizak, V., & Song, D. D. (2005). The impact of irrelevant dimensional variation on rule-based category learning in patients with Parkinson's disease. *Journal of the International Neuropsychological Society, 11*(5), 503-513.

Filoteo, J. V., Maddox, W. T., Salmon, D. P., & Song, D. D. (2005). Information-integration category learning in patients with striatal dysfunction. *Neuropsychology, 19*(2), 212-222.

Filoteo, J. V., Maddox, W. T., Simmons, A. N., Ing, A. D., Cagigas, X. E., Matthews, S., & Paulus, M. P. (2005). Cortical and subcortical brain regions involved in rule-based category learning. *Neuroreport, 16*(2), 111-115.

Filoteo, J. V., Maddox, W. T., Song, D., & Salmon, D. P. (2007). Implicit category learning performance predicts rate of cognitive decline in nondemented patients with Parkinson's disease. *Neuropsychology*,

Filoteo, J. V., Paul, E. J., Ashby, F. G., Frank, G. K.W., Helie, S., Rockwell, R., Bischoff-Grethe, A., Wierenga, C., & Kaye, W. H. (2014). Simulating category learning and set shifting deficits in patients weight-restored from Anorexia Nervosa. *Neuropsychology*, *28*, 741-751.

Filoteo, J. V., Simmons, A. N., Zeithamova, D., Maddox, W. T., & Paulus, M. P. (2006). *Change in patterns of brain activity related to early and later learning of information-integration category structures.* Paper presented at the Cognitive Neuroscience Society, San Francisco.

Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., & Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *NeuroImage*, *32*, 956-967.

Frank, G. K., Bailer, U. F., Henry, S. E., Drevets, W., Meltzer, C. C., Price, J. C., . . . Kaye, W. H. (2005). Increased dopamine D2/D3 receptor binding after recovery from anorexia nervosa measured by positron emission tomography and [11C] raclopride. *Biological Psychiatry*, 58, 908–912.

Gilbert, B., Belleville, S., Bherer, L., & Chouinard, S. (2005). Study of verbal working memory in patients with Parkinson's disease. *Neuropsychology, 19*(1), 106-114.

Gluck, M. A., Oliver, L. M., & Myers, C. E. (1996). Late-training amnesic deficits in probabilistic category learning: a neurocomputational analysis. *Learning & Memory*, *3*(4), 326-340.

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: individual variability in strategies for probabilistic category learning. *Learing & Memory, 9*(6), 408-418.

Harrison A, O'Brien N, Lopez C, & Treasure J. (2010). Sensitivity to reward and punishment in eating disorders. *Psychiatry Research*. 177:1–11.

Harrison A, Treasure J, & Smillie LD. (2011). Approach and avoidance motivation in eating disorders. *Psychiatry Research*. 188:396–40

Heaton, R. K. (1981). *Wisconsin Card Sorting Test*. Odessa, FL: Psychological Assessment Resources.

Hélie, S., Paul, E. J., & Ashby, F. G. (2012a). A neurocomputational account of cognitive deficits in Parkinson's disease. *Neuropsychologia*, *50*, 2290-2302.

Hélie, S., Paul, E. J., & Ashby, F. G. (2012b). Simulating the effects of dopamine imbalance on cognition: From positive affect to Parkinson's disease. *Neural Networks*, *32*, 74-85.

Jappe, L. M., Frank, G. K. W., Shott, M. E., Rollin, M. D. H., Pryor, T., Hagman, J. O., . . . Davis, E. (2011). Heightened sensitivity to reward and punishment in anorexia nervosa. *International Journal of Eating Disorders*, 44, 317–324.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*(5280), 1399-1402.

Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: parallel brain systems for item memory and category knowledge. *Science, 262*(5140), 1747-1749.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

Maddox, W.T. (1999). On the dangers of averaging across subjects when comparing decision bound and generalized context models of categorization. *Perception and Psychophysics*, 61, 354–374.

Maddox, W. T., Aparicio, P., Marchant, N. L., & Ivry, R. B. (2005). Rule-based category learning is impaired in patients with Parkinson's disease but not in patients with cerebellar disorders. *Journal of Cognitive Neuroscience, 17*(5), 707-723.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics, 53*, 49-70.

Maddox, W. T., Bohil, C. J., & Ing, A. D. (2004). Evidence for a procedural-learning-based system in perceptual category learning. *Psychonomic Bulletin & Review, 11*, 945-952.

Maddox, W. T., & Filoteo, J. V. (2001). Striatal contributions to category learning: quantitative modeling of simple linear and complex nonlinear rule learning in patients with Parkinson's disease. *Journal of the International Neuropsychological Society, 7*(6), 710-727.

Maddox, W. T., Filoteo, J. V., Delis, D. C., & Salmon, D. P. (1996). Visual selective attention deficits in patients with Parkinson's disease: A quantitative model-based approach. *Neuropsychology, 10*(2), 197-218.

Maddox, W. T., Filoteo, J. V., Hejl, K. D., & Ing, A. D. (2004). Category number impacts rule-based but not information-integration category learning: further evidence for dissociable category-learning systems. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*(1), 227-245.

Maddox, W. T., Filoteo, J. V., & Lauritzen, J. S. (2007). Within-category discontinuity interacts with verbal rule complexity in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 33*(1), 197-218.

Maddox, W. T., Filoteo, J. V., Lauritzen, J. S., Connally, E., & Hejl, K. D. (2005). Discontinuous categories affect information-integration but not rule-based category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 31*(4), 654-669.

Maddox, W. T., Filoteo, J. V., & Zeithamova, D. (2010). Computational models inform clinical science and assessment: An application to category learning in striatal-damaged patients. *Journal of Mathematical Psychology*, 54, 109-122.

Mattis, S. (1988). *Dementia Rating Scale*. Odessa, FL: Psychological Assessment Resources.

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*(2), 438-442.

Nomura, E. M., Maddox, W. T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., Mesulam, M. M., & Reber, P. J. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex, 17*(1), 37-43.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101*(1), 53-79.

Owen, A. M., Beksinska, M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., Sahakian, B. J., & Robbins, T. W. (1993). Visuospatial memory deficits at different stages of Parkinson's disease. *Neuropsychologia, 31*(7), 627-644.

Owen, A. M., Iddon, J. L., Hodges, J. R., Summers, B. A., & Robbins, T. W. (1997). Spatial and non-spatial working memory at different stages of Parkinson's disease. *Neuropsychologia, 35*(4), 519-532.

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia, 41*, 245-251.

Postle, B. R., Jonides, J., Smith, E. E., Corkin, S., & Growdon, J. H. (1997). Spatial, but not object, delayed response is impaired in early Parkinson's disease. *Neuropsychology, 11*(2), 171-179.

Price, A. L. (2005). Cortico-striatal contributions to category learning: dissociating the verbal and implicit systems. *Behavioral Neuroscience, 119*(6), 1438-1447.

Reber, P. J., & Squire, L. R. (1999). Intact learning of artificial grammars and intact category learning by patients with Parkinson's disease. *Behavioral Neuroscience, 113*(2), 235-242.

Reber, P. J., Stark, C. E., & Squire, L. R. (1998). Contrasting cortical activity associated with category memory and recognition memory. *Learning & Memory, 5*(6), 420-428.

Roberts, M. E., Tchanturia, K., Stahl, D., Southgate, L., & Treasure, J. (2007). A systematic review and meta-analysis of set-shifting ability in eating disorders. *Psychological Medicine*, 37, 1075–1084.

Roberts, M. E., Tchanturia, K., & Treasure, J. L. (2010). Exploring the neurocognitive signature of poor set-

shifting in anorexia and bulimia nervosa. *Journal of Psychiatric Research*, 44, 964–970.

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain, 127*(Pt 4), 851-859.

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*(4), 676-686.

Shott, M. E., Filoteo, J. V., Bhatnagar, K. A., Peak, N. J., Hagman, J. O., Rockwell, R., . . . Frank, G. K. W. (2012). Cognitive set-shifting in anorexia nervosa. *European Eating Disorders Review*, 20, 343–349.

Skosnik, P. D., Mirza, F., Gitelman, D. R., Parrish, T. B., Mesulam, M-M, & Reber, P. J. (2002). Neural correlates of artificial grammar learning. *NeuroImage, 17*, 1306-1314.

Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition, 65*(2-3), 167-196.

Smith, E. E., & Sloman, S. A. (1994). Similarity- versus rule-based categorization. *Memory & Cognition, 22*(4), 377-386.

Steinglass, J. E., Walsh, B. T., & Stern, Y. (2006). Set shifting deficit in anorexia nervosa. Journal of the International Neuropsychological Society, 12, 431–435.Vakil, E., & Herishanu-Naaman, S. (1998). Declarative and procedural learning in Parkinson's disease patients having tremor or bradykinesia as the predominant symptom. *Cortex, 34*(4), 611-620.

Tchanturia, K., Harrison, A., Davies, H., Roberts, M., Oldershaw, A., Nakazato, M., . . . Treasure, J. (2011). Cognitive flexibility and clinical severity in eating disorders. *PLoS ONE*, 6, e20462

Tchanturia, K., Morris, R. G., Surguladze, S., & Treasure, J. (2002). An examination of perceptual and cognitive set shifting tasks in acute anorexia nervosa and following recovery. *Eating and Weight Disorders*, 7, 312–315.

Waldschmidt, J. G., & Ashby, F. G. (2011). Cortical and striatal contributions to automaticity in information-integration categorization. *Neuroimage, 56*, 1791-1802.

Wickens, T. D. (2002). *Elementary signal detection theory*. Oxford University Press, USA.

Witt, K., Nuhsman, A., & Deuschl, G. (2002). Dissociation of habit-learning in Parkinson's and cerebellar disease. *Journal of Cognitive Neuroscience, 14*(3), 493-499.

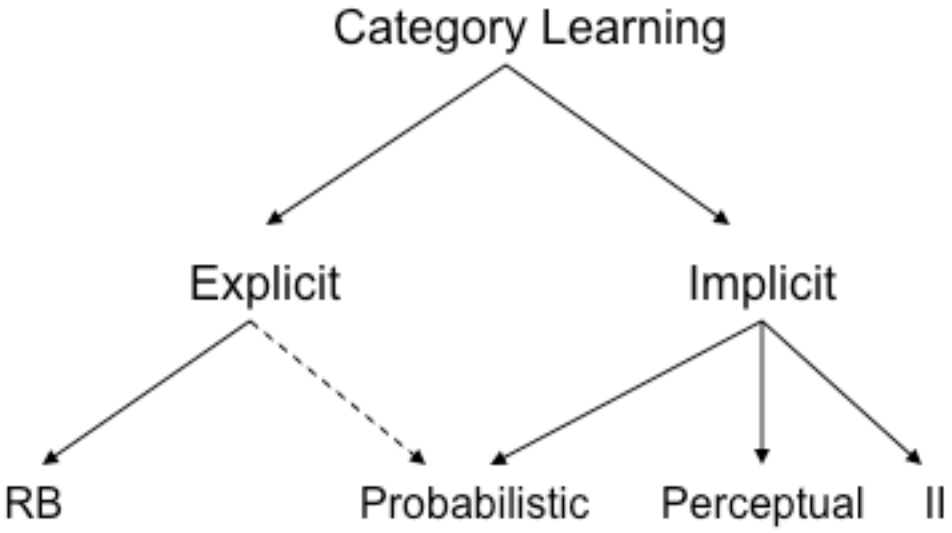

Figure 1. Classification of explicit and implicit category learning tasks.

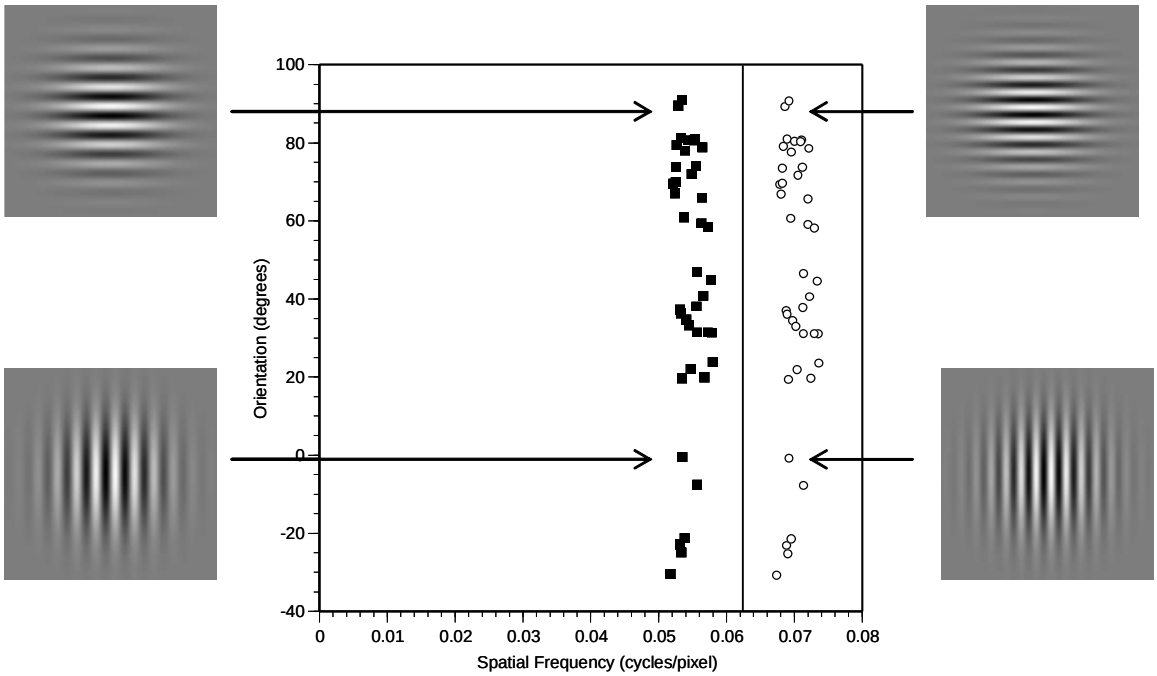

Figure 2. Stimulus distributions and sample stimuli used in a rule-based (RB) perceptual categorization task. Filled squares represent stimuli from Category A and open circles represent stimuli from Category B. The solid line represents the optimal one-dimensional decision bound. Arrows point from specific stimulus exemplars to their location in the two-dimensional stimulus space. From Filoteo, Maddox, Ing, and Song (2007).

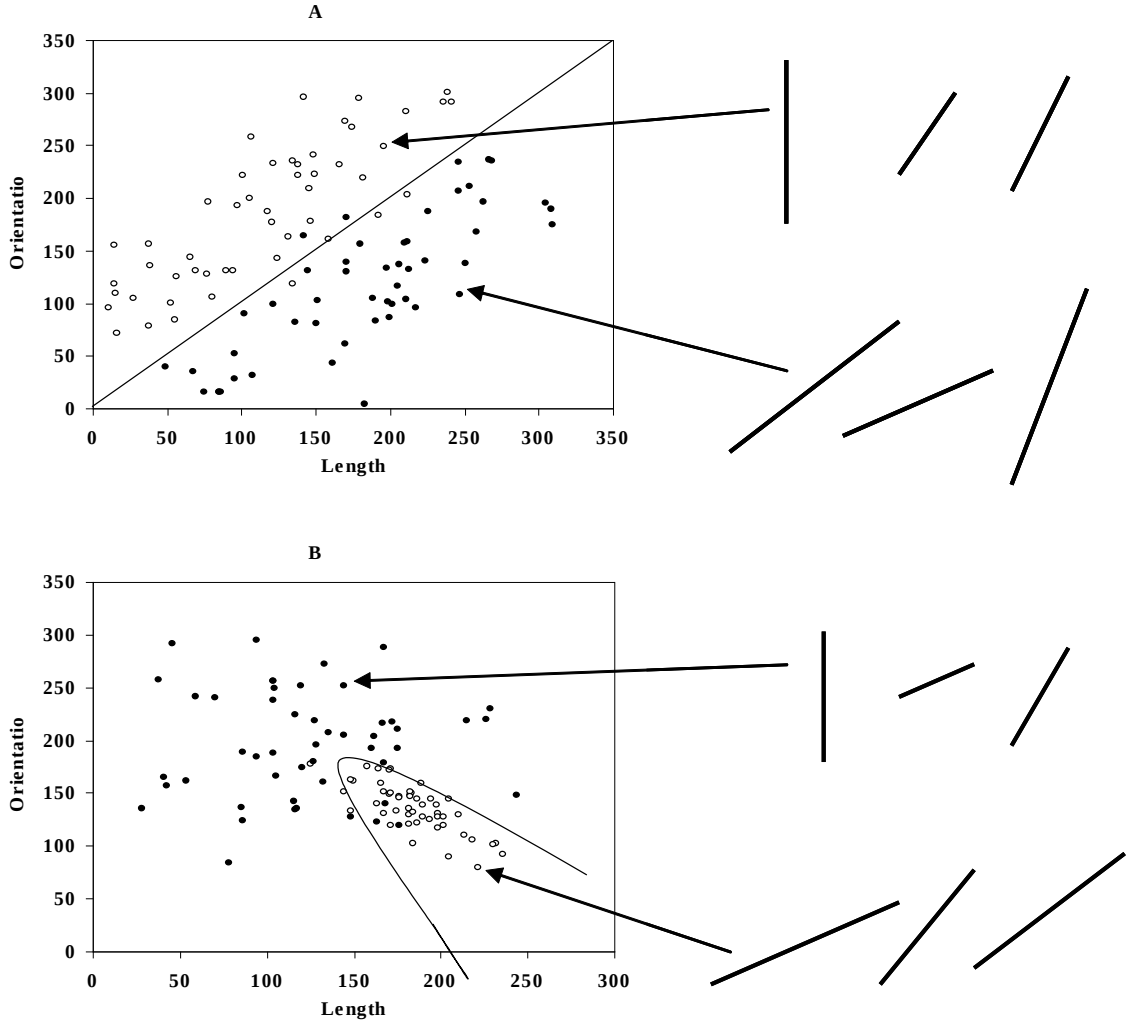

Figure 3. Sample stimuli and stimulus distributions for (A) the linear information-integration (II) task and (B) the nonlinear II task in which the stimuli were single lines that vary in length and orientation. Open circles represent stimuli from Category A and closed circles represent stimuli from Category B. The solid line and curve represent the optimal rules. Arrows point from specific stimulus exemplars to their location in the two-dimensional stimulus space. From Filoteo, Maddox, Salmon, and Song (2005).

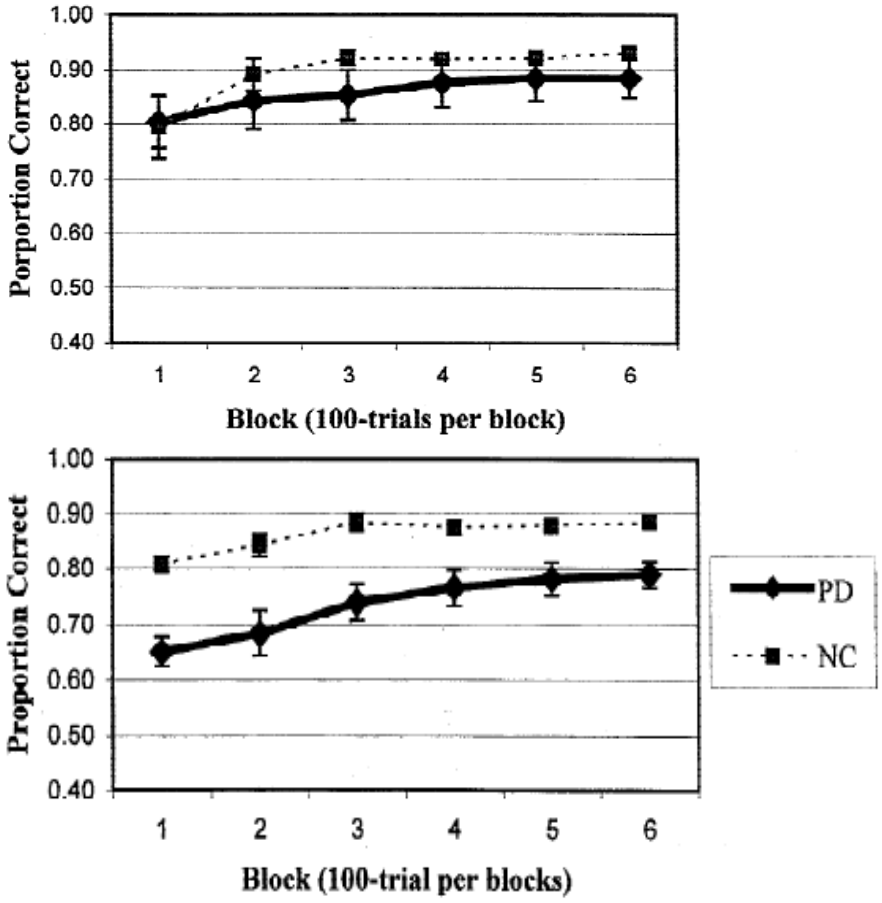

Figure 4. Accuracy (proportion correct) for the patients with Parkinson's disease (PD) and normal controls (NC) for the 600-trial session in 100-trial blocks for (a) the linear II task, and (b) the nonlinear II task. Standard errors are also included. From Maddox & Filoteo (2001).

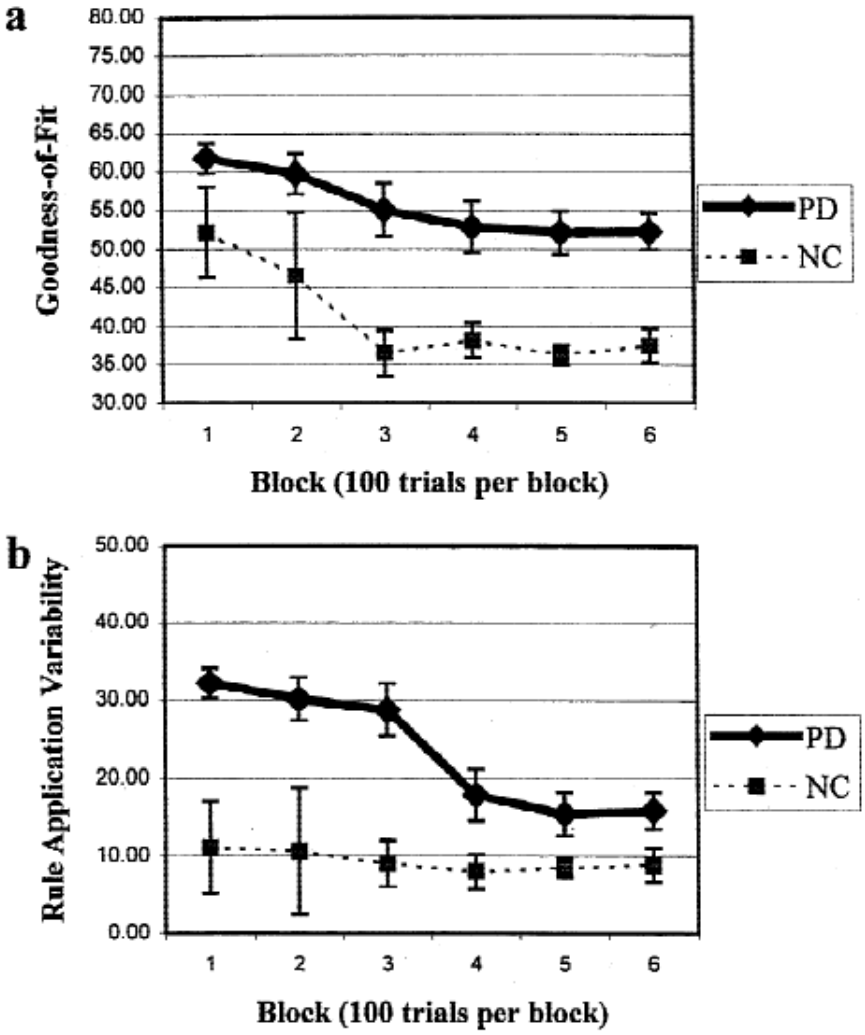

Figure 5. (a) Fit values, and (b) application variability estimates for the patients with Parkinson's disease (PD) and normal controls (NC) for the 600 trial experimental session in 100-trial blocks in the nonlinear integration experiment. From Maddox & Filoteo (2001).

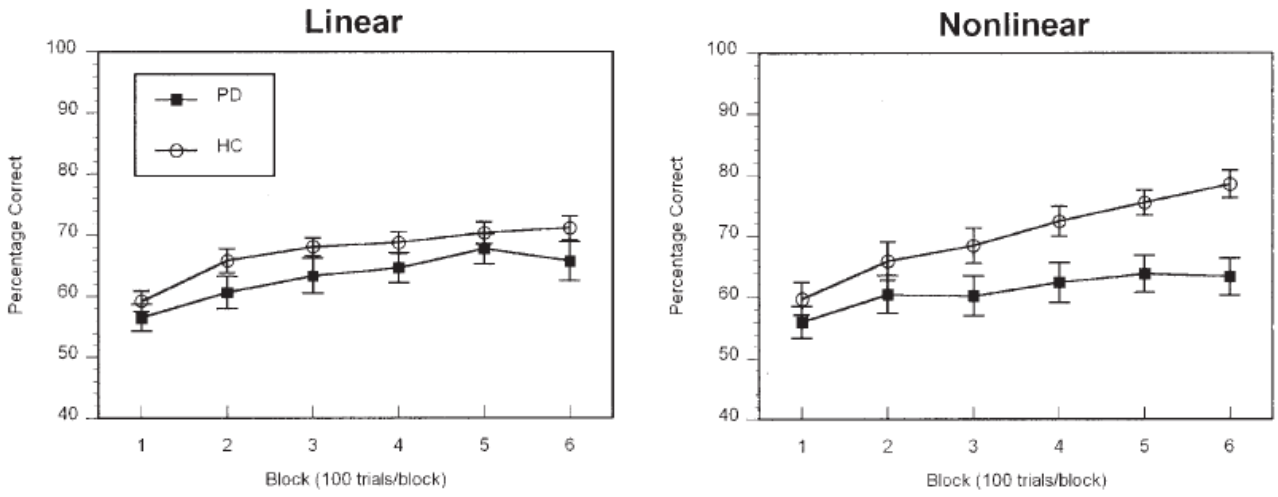

Figure 6. Overall accuracy for linear and nonlinear conditions for patients with Parkinson's disease (PD) and Healthy Control (HC) participants. Error bars denote standard error of the mean. From Filoteo, Maddox, Salmon et al. (2005).

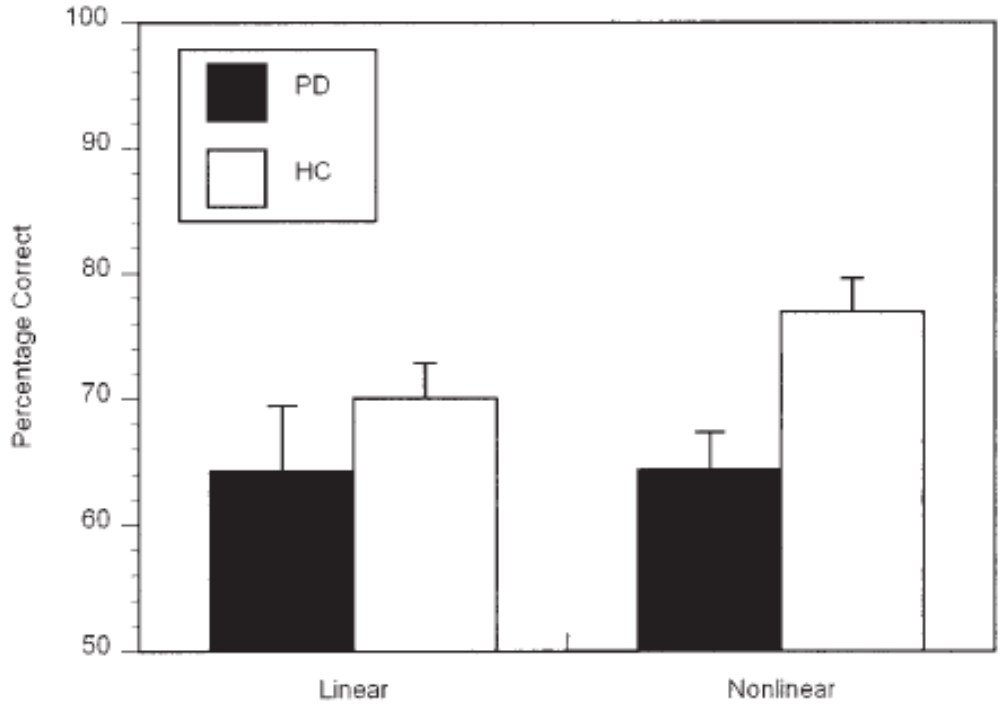

Figure 7. Accuracy rates for Parkinson's disease (PD) patients and healthy control (HC) participants whose data were best fit by a procedural-learning model in the last block of trials in the linear and nonlinear II tasks. Error bars denote standard error of the mean. From Filoteo, Maddox, Salmon et al. (2005).

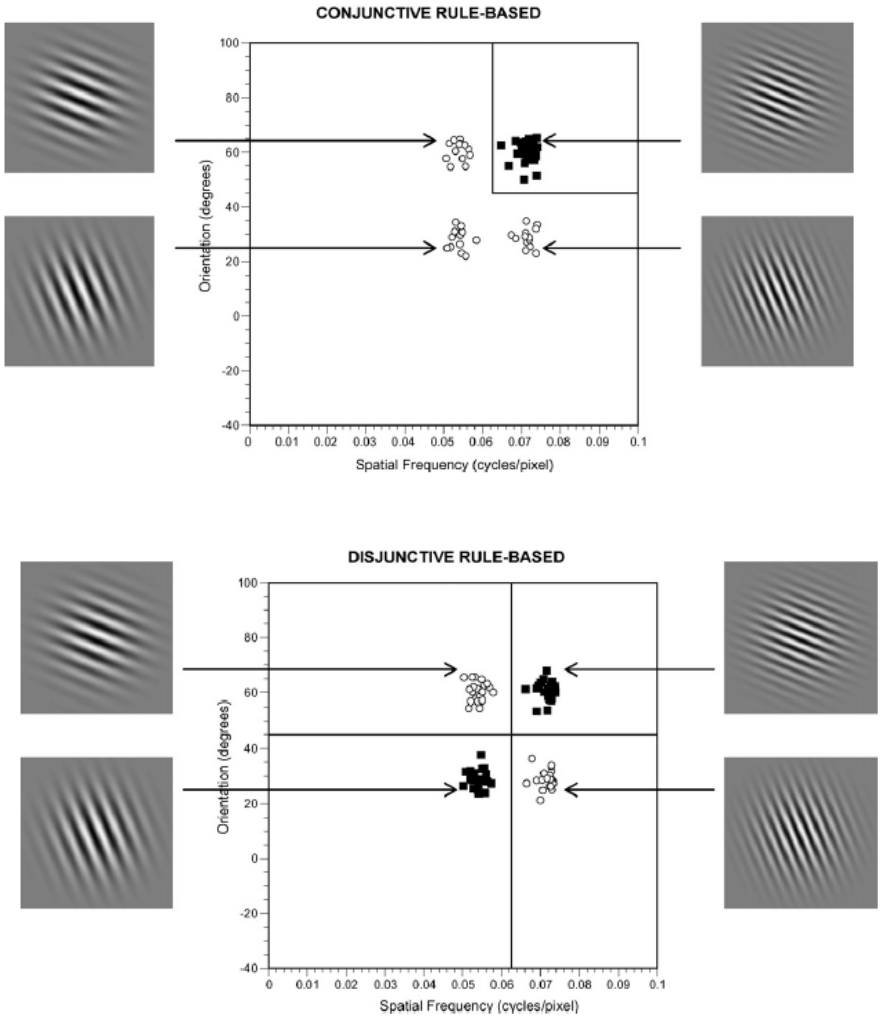

Figure 8. Stimulus distributions for (A) conjunctive, and (B) disjunctive rule-based category learning tasks. Filled squares represent stimuli from Category A and open circles represent stimuli from Category B. Solid lines represent the optimal bounds for each condition. From Filoteo, Maddox, Ing, and Song (2007).

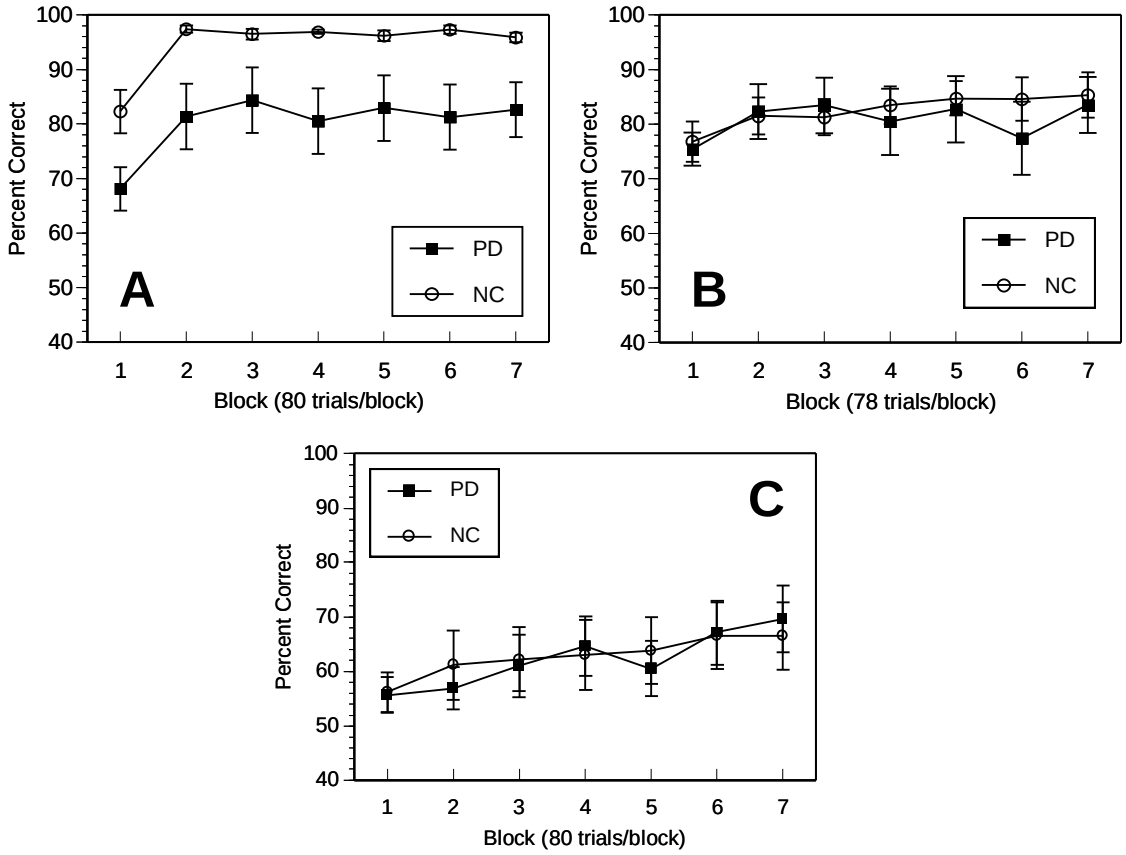

Figure 9. Accuracy for Parkinson's disease (PD) patients and normal control (NC) participants for (A) one-dimensional, (B) conjunctive, and (C) disjunctive rule-based category learning conditions. From Filoteo, Maddox, Ing, and Song (2007).

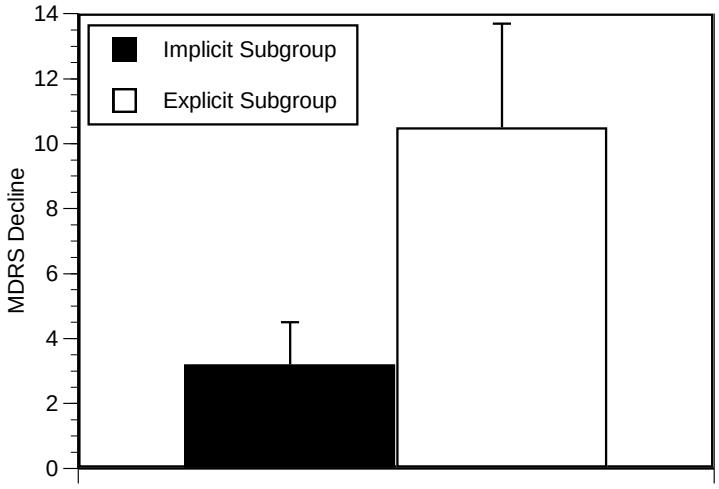

Figure 10. Decline on the Mattis Dementia Rating Scale in Parkinson's disease patient subgroups whose data were best fit by an implicit model or an explicit model.

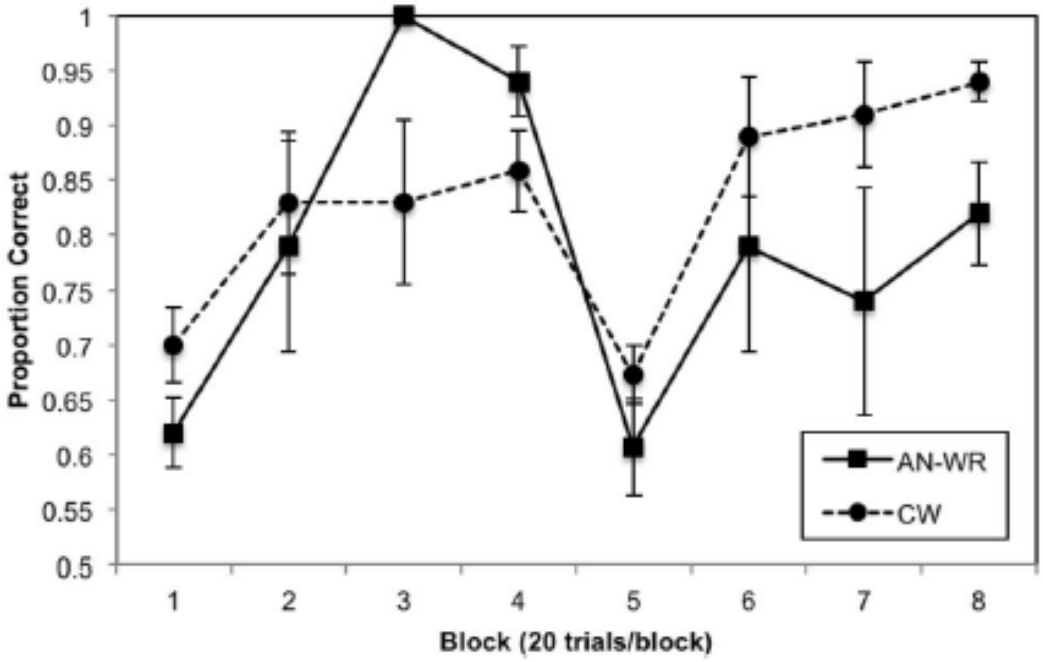

Figure 11. Accuracy (proportion correct) for weight restored Anorexic patients (AN-WR) and control women (CW) groups (error bars are standard error of the mean) in the study by Filoteo et al. (2014).