Neuropsychology
2014, Vol. 28, No. 6, 945–958

0894-4105/14/$12.00 http://dx.doi.org/10.1037/neu0000121

A Meta-Analysis and Meta-regression of Serial Reaction Time Task Performance in Parkinson's Disease

Gillian M. Clark and Jarrad A. G. Lum
Deakin University

Michael T. Ullman
Georgetown University

Objective: This article reports findings of a meta-analysis and meta-regression summarizing research on implicit sequence learning in individuals with Parkinson's disease (PD), as measured by the Serial Reaction Time (SRT) task. ***Method:*** Following a systematic search of the literature, we analyzed a total of 27 studies, representing data from 505 participants with PD and 460 neurologically intact control participants. ***Results:*** Overall, the meta-analysis indicated significantly ($p < .001$) worse sequence learning by the PD group than the control group. The average weighted effect size was found to be .531 (95% CI [.332, .470]), which is a medium effect size. However, moderate to high levels of heterogeneity (differences) were found between study effect sizes ($I^2 = 58\%$). Meta-regression analysis suggested that presentation of the SRT task under dual task conditions coupled with PD severity or characteristics of the sequence might affect study effect sizes. ***Conclusions:*** The meta-analysis provides clear support that learning in procedural memory (procedural learning), which underlies implicit sequence learning in the SRT task, is impaired in PD.

Keywords: serial reaction time (RT), implicit learning, procedural learning, Parkinson's disease, meta-analysis

Supplemental materials: http://dx.doi.org/10.1037/neu0000121.supp

Parkinson's disease (PD) is a neurodegenerative disorder associated with the death of dopamine-producing cells in the substantia nigra, leading to the dysfunction of the basal ganglia and ensuing motor symptoms (Dubois & Pillon, 1996; Lang & Lozano, 1998). These symptoms include tremor, rigidity, bradykinesia, and postural instability (Jankovic, 2008; Lozano et al., 1995). Research has also investigated the extent to which PD is associated with cognitive impairments (Cooper, Sagar, Jordan, Harvey, & Sullivan, 1991; Dubois & Pillon, 1996; Owen, Iddon, Hodges, Summers, & Robbins, 1997). Within this literature, numerous studies have investigated procedural memory in PD, in particular with Serial Reaction Time (SRT) tasks (e.g., Gawrys et al., 2008; Muslimovic, Post, Speelman, & Schmand, 2007; Price & Shin, 2009; Seidler, Tuite, & Ashe, 2007; van Tilborg & Hulstijn, 2010; Vandenbossche et al., 2013; Wang, Sun, & Ding, 2009). This report uses meta-analysis and meta-regression to summarize this research, updating and extending a previous meta-analysis on this topic (Siegert, Taylor, Weatherall, & Abernethy, 2006).

This article was published Online First July 7, 2014.

Gillian M. Clark and Jarrad A. G. Lum, School of Psychology, Deakin University; Michael T. Ullman, Department of Neuroscience, Georgetown University.

Correspondence concerning this article should be addressed to Jarrad A. G. Lum, School of Psychology, Deakin University, 221 Burwood Highway, Burwood, Victoria, Australia, 3121. E-mail: jarrad.lum@deakin.edu.au

Procedural Memory

It is widely accepted that there are multiple memory systems in the brain (Knowlton, Mangels, & Squire, 1996; Squire, 2004; Squire, Knowlton, & Musen, 1993; Squire & Zola, 1996). The procedural memory system is one of a number of nondeclarative memory systems that supports the learning and access of implicit (nonconscious) knowledge (Knowlton et al., 1996; Knowlton, Squire, & Gluck, 1994; Ullman, 2001; Yin & Knowlton, 2006). Learning in the procedural system is gradual, with repetition or practice required in order for skills or knowledge to be acquired. However, once learnt, knowledge can be accessed rapidly and without awareness. Much is known about the neural substrates of the procedural memory system. Evidence from both healthy individuals and those with neurological dysfunctions has repeatedly implicated the basal ganglia and motor-related and prefrontal areas, as well as the cerebellum (Kandel, Schwartz, & Jessell, 2012; Packard & Knowlton, 2002; Parent & Hazrati, 1995; Pascual-Leone, Wassermann, Grafman, & Hallett, 1996; Ullman, 2004, 2006).

The Serial Reaction Time Task

The SRT task developed by Nissen and Bullemer (1987) has been used to investigate learning in procedural memory (i.e., procedural learning) in a wide range of nonclinical (e.g., Lum & Kidd, 2012; Thomas et al., 2004; Thomas & Nelson, 2001) and clinical populations (e.g., Ferraro, Balota, & Connor, 1993; Knopman & Nissen, 1991; Siegert, Weatherall, & Bell, 2008), including individuals with PD (e.g., Siegert et al., 2006; Stefanova, Kostic, Ziropadja, Markovic, & Ocic, 2000; van Tilborg & Hulstijn, 2010;

Vandenbossche et al., 2013; Wang et al., 2009; Werheid, Ziessler, Nattkemper, & Von Cramon, 2003; Werheid, Zysset, Muller, Reuter, & Von Cramon, 2003; Westwater, McDowall, Siegert, Mossman, & Abernethy, 1998).

The standard protocol for the SRT task requires participants to be seated in front of a computer display that shows a visual stimulus repeatedly appearing in one of four locations (for an overview of SRT task methodology see Robertson, 2007). Stimulus presentations are usually grouped into blocks typically comprising around 80 to 100 stimuli, although this figure varies substantially between studies (see Lum, Ullman, & Conti-Ramsden, 2013). On most blocks, stimulus presentations follow a predefined sequence. Depending on the study, the length of the sequence may vary from 8 to 12. In the implicit version of the SRT task, which is examined in this report, participants are not informed that the visual stimulus follows a sequence. After a series of "sequenced blocks" a "random block" is presented in which the visual stimulus appears in a random order. The only instruction provided to participants, at the beginning of the task, is to indicate the location of the visual stimulus on each trial. In most studies participants respond to the stimulus by pressing one of several buttons on a response box (e.g., Pascual-Leone et al., 1993; Shin & Ivry, 2003). However, in other work the motor demands of the task have been reduced by asking participants to provide a verbal response (e.g., Smith & McDowall, 2004; Smith, Siegert, McDowall, & Abernethy, 2001).

The key dependent variable in SRT tasks is reaction time (RT), which measures how fast participants identify the location of the visual stimulus. In neurologically intact individuals, RTs gradually decrease (i.e., responses become faster) across the series of "sequenced blocks." Then, during the "random block" RTs increase again (e.g., Lum & Kidd, 2012; Thomas et al., 2004; Thomas & Nelson, 2001). This RT increase from the sequence block to the random block is taken to suggest that knowledge or information about the sequence has been learnt. Importantly, the change in RTs is observed even though participants are not able to explicitly recall the sequence. Neuroanatomical meta-analysis of the functional neuroimaging literature examining the SRT task indicates that the task indeed depends on the neural substrates of procedural memory, activating both the basal ganglia (i.e., the putamen) and the cerebellum (Hardwick, Rottschy, Miall, & Eickhoff, 2013).

SRT Task Performance in Parkinson's Disease

The performance of individuals with PD on the SRT task provides an important opportunity to examine the status of sequence learning in procedural memory in the disorder. Given that learning in the SRT task appears to depend on the procedural memory system, including the basal ganglia, individuals with PD may be expected to perform worse on the task than healthy comparison groups. Specifically, the increase in RTs from the final sequence block to the following random block should be smaller in PD groups relative to healthy controls. That is, a significant Group (PD vs. Control) × Block (Sequence vs. Random) interaction is expected.

A meta-analysis that summarized research published between 1987 and 2005 on SRT task performance in PD was presented by Siegert, Taylor, Weatherall, and Abernethy (2006). In typical meta-analyses, effect sizes and variances from studies with similar methodologies are pooled and an average effect size is computed (Borenstein, Hedges, Higgins, & Rothstein, 2011; Hunter, Schmidt, & Jackson, 1982). The meta-analysis undertaken by Siegert et al. (2006) summarized the results from six studies with a combined sample size of 67 individuals with PD and 87 neurologically intact controls. The average effect size computed in the meta-analysis was for the Group (PD vs. Control) × Block (Sequence vs. Random) interaction. Using a random effects model to average study results, the average effect size was found to be 0.65 (95% CI [0.10, 1.20]), and was statistically significant. This result indicates that on average, the increase in RTs from the sequence block to the random block was 0.65 standard deviations larger in the control groups than the PD groups. This corresponds to a medium effect size according to Cohen's (1988) taxonomy. Siegert et al.'s (2006) meta-analysis also identified considerable levels of heterogeneity, that is, differences between individual study effect sizes. Quantification of the variability in effect sizes using the I^2 statistic indicated that 64.8% of the heterogeneity could not be explained by random error or chance. This suggests there may be one or more systematic influences on study findings. Thus an outstanding question arising from Siegert et al.'s meta-analysis is, what variable or variables assert a systematic influence on study findings?

Inspection of the PD/SRT task literature reveals a number of potential candidate variables that may account for variability in study effect sizes. First, the average severity of the PD group's symptoms might account for differences between study findings. Price and Shin (2009) found that participants with moderate PD symptoms, but not those whose symptoms were mild, performed significantly worse on their version of the SRT task than controls. Inspection of the PD/SRT task literature reveals variability in the average severity of PD symptoms when quantified using the Hoehn-Yahr scale (Hoehn & Yahr, 1967), which measures the severity of PD symptoms on a 5-point scale, with lower values corresponding to milder symptoms. In some studies the average Hoehn-Yahr rating was approximately 1.5 (e.g., Stefanova et al., 2000; van Tilborg & Hulstijn, 2010; Werheid, Zysset, et al., 2003; Westwater et al., 1998), yet in others the average severity was about 3 (e.g., Deroost, Kerckhofs, Coene, Wijnants, & Soetens, 2006; Price & Shin, 2009). Studies that have PD participants with more severe symptoms might have observed larger sequence-random differences on the SRT task between PD and control groups.

Second, differences in the input method used to collect responses on the SRT task may systematically influence study results. The most common method in SRT studies of PD uses a response box (e.g., Sommer, Grafman, Clark, & Hallett, 1999; Stefanova et al., 2000; van Tilborg & Hulstijn, 2010). However, this method may disproportionally disadvantage PD participants because a central feature of the disorder is motor problems. Thus, the extent to which the SRT task places demands on motor skills may contribute to whether a study observes a significant difference between groups. To attempt to address this issue some investigators have modified the SRT task so that participants indicate the location of the visual stimulus with a verbal response (Smith & McDowall, 2004; Sommer et al., 1999; Stefanova et al., 2000; van Tilborg & Hulstijn, 2010). Using this method, nonsignificant differences between PD and control groups have indeed been reported (Smith et al., 2001). If the method used to collect responses

influences study findings, smaller effect sizes may be observed for studies in which subjects respond verbally.

Third, presentation of the SRT task under dual task conditions may also explain differences between study findings. SRT dual task paradigms require participants to engage in a second activity while simultaneously responding to the visual stimulus. In the PD literature, participants are usually asked to count tones while completing the SRT task (Kelly, Jahanshahi, & Dirnberger, 2004; Seidler et al., 2007; Vandenbossche et al., 2013). This approach is intended to reduce the possibility that participants will gain explicit awareness of the sequence, in order to encourage learning in procedural memory. However, research has also shown that dual task performance of any sort is disproportionately poorer in PD groups compared with control groups (Dalrymple-Alford, Kalders, Jones, & Watson, 1994; Wu & Hallett, 2008). Thus, studies using a dual task paradigm to study SRT task performance in PD may observe larger differences between study and control groups.

Fourth, the type of sequence used in SRT tasks might also influence outcomes in studies of PD. Sequences in SRT tasks can vary regarding the extent to which elements in a given location are first-order conditional (FOC) or second-order conditional (SOC). In FOC sequences, each element in the sequence can be at least partially predicted from the preceding element. For example, in the sequence 4243123142, the item 1 is always followed by either a 2 or a 4 (50% probability each), but never a 3 (0%); similarly the item 3 is always followed by a 1 (100% probability), but never by a 2 or 4 (0%). In contrast, for SOC sequences the probability between element transitions is equal. For example, in the sequence 134231432412, there is a 33.3% probability that 1 will be followed by 2, 3, and 4 (Robertson, 2007). In a SOC sequence, all transitions between elements can be considered ambiguous (Cohen, Ivry, & Keele, 1990).

The sequences that have been used to investigate implicit learning in PD have varied with respect to the number of ambiguous element transitions, that is, the number of SOC elements. For example, DeRoost, Kerckhofs, Coene, Wijnants, and Soetens (2006) presented a fully SOC sequence to PD and control groups. In this study a 12-element sequence, 121342314324, was used for the SRT task. In the sequence there are a total of 12 transitions (because the sequence is repeated, the last element in the sequence is followed by the first). The probability of one element following another is 33.3% for all transitions. Thus, in DeRoost et al. (2006), 12 transitions out of a possible 12 can be considered to be ambiguous. In contrast, Smith and McDowall (2006) used an 8-item sequence, 14213243, where none of the transitions were ambiguous. Rather, each element in the sequence can be predicted, to some extent, from the preceding element. For example, in this sequence the case of a 1 being followed by 2 occurs 0% of the time and the probability of 3 or 4 following 1 is 50% each.

The number of ambiguous transitions between elements may have an impact on study findings. Results from several studies indicate that the implicit learning of SOC sequences may depend on medial temporal lobe structures (Curran, 1997; Ergorul & Eichenbaum, 2006; Schendan, Searl, Melrose, & Stern, 2003). One explanation offered to account for this finding is that learning ambiguous transitions between elements requires being able to bind items that are arbitrarily related, a type of learning that is well support by the medial temporal lobes (Mayes, Montaldi, & Migo, 2007; Poldrack & Rodriguez, 2003; Robertson, 2007; Squire, Stark, & Clark, 2004). As the learning and memory functions that are supported by the medial temporal lobes appear relatively spared in PD (Helkala, Laulumaa, Soininen, & Riekkinen, 1988; Ullman et al., 1997), smaller differences between study and comparison groups might be observed in studies that use sequences with ambiguous transitions.

Fifth, some evidence suggests that the length of the sequence may systematically influence study results. Pascual-Leone et al. (1993) found that the magnitude of the difference between PD and control groups was larger when participants were presented with a 12-element FOC sequence compared to an 8-element FOC sequence. SRT tasks using sequences that are short in length may place fewer demands on procedural memory, leading to increased learning.

Finally, evidence also suggests that the number of times the sequence is presented may influence results. In a meta-analysis investigating SRT task performance in the neurodevelopmental disorder of dyslexia, Lum, Ullman, and Conti-Ramsden (2013) found smaller differences between dyslexia and control groups for those studies that presented the sequence more times. It was suggested that more exposures to the sequence might increase the likelihood that the sequence will be acquired by the procedural memory system, even when the system is dysfunctional.

Aims of the Current Report

The aim of this report was, first of all, to update and expand the meta-analysis examining SRT task performance in PD undertaken by Siegert et al. (2006), which summarized the results of six studies. In the current meta-analysis, we included results from 27 studies, representing data from 505 participants with PD and 460 neurologically intact control participants. The second aim of the study was to investigate variables that may account for differences in study findings. Specifically, using meta-regression we investigated whether individual study results could be predicted by one or more of the following variables or their interactions: the participant-level factor of PD severity, and several SRT task variables, namely the type of input device (response method) that was used, whether the task was administered under single task or dual task conditions, the sequence type, the sequence length, and the number of exposures to the sequence.

Method

Study Design

Electronic databases of published and unpublished/gray literature were searched. A search for published literature was undertaken using Psychological Information Database (PsycINFO), Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medical Database (EMBASE), Cumulative Index to Nursing and Allied Health Literature (CINAHL), and PubMed. Unpublished literature was searched using BioSciences Information Service of Biological Abstracts (BIOSIS), OpenGrey, ProQuest Dissertation and Conference Abstracts, and PsycExtra. Details of all keywords, fields searched, Boolean operators, and syntax are presented in the online supplemental materials. The search strategy for published studies was executed in April, 2013.

The search strategy for unpublished studies was executed in January, 2014.

Study Inclusion Criteria

Following execution of the search, articles were only included in the meta-analysis if they met the following inclusionary criteria, which were similar to those used by Siegert et al. (2006). First, only studies involving human participants were included. It was further required that the SRT task was presented to one group of participants comprising individuals with PD and a control group comprising neurologically intact individuals of comparable age. Second, because the SRT task was first described in an article published in 1987 (i.e., Nissen & Bullemer, 1987), studies were excluded if they were published prior to this year. Third, the study was required to report original research. Fourth, the study needed to have administered a version of Nissen and Bullemer's SRT task. That is, the study was required to have presented an implicit version of the task (i.e., participants were not informed of the sequence). Studies that presented an explicit version of the SRT task were not included in the review. Also, the structure of the task must have presented a series of sequence blocks followed by a random block. Finally, to ensure testing conditions between studies were similar, we excluded one study where participants underwent transcranial magnetic stimulation while completing the task (Pascual-Leone et al., 1994). Figure 1 summarizes studies removed following application of each criterion according to PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009).

Study Selection

After the removal of duplicates, one of the authors (GC) assessed all the abstracts. A second author (JL) assessed a random sample of 10% of all abstracts. The two authors independently screened full-text articles according to the eligibility criteria described above. There was 100% agreement on these articles. A total of 27 studies were included, and their data extracted for meta-analysis. A summary of each study's participants and characteristics of the SRT task structure is presented in Tables 1 and 2, respectively.

Values for the participant and task characteristics shown in Tables 1 and 2 were extracted directly from the studies, except for the proportion of ambiguous transitions presented in Table 2, which was obtained by extracting the sequence used in each study. From this information the proportion of ambiguous transitions was calculated by dividing the number of ambiguous transitions into the total number of transitions in the sequence. A transition was considered ambiguous when the occurrence of a single element gave no information about the following element. For example, the transitions 1–2, 1–3, and 1–4 were considered as three ambiguous transitions if they each occurred once within a sequence. However, if only the transitions 1–2 and 1–3 were present (and not 1–4), these were not considered ambiguous transitions.

Effect Size Calculations and Data Extraction Procedures

The widely accepted method adopted to compare performance on the SRT task between two groups is to calculate whether the difference in RTs between the random block and the preceding sequence block differs between a control and study group. That is, whether a significant Group × Block interaction is present. In undertaking this meta-analysis, results reported in individual studies were extracted so that the effect size for the interaction and variance could be computed. Following the method by Siegert et al. (2006), the effect size measure used in this study was the standardized mean difference also known as Cohen's *d*. This describes the difference between groups on the SRT task in standard deviation units. Cohen's *d* is known to have a slight bias when studies have small sample sizes, and so a correction factor was applied that reduces this bias. Cohen's *d* was calculated so that positive values indicated that the control group in each study displayed higher levels of procedural learning than the group with PD or alternatively, the PD group performed worse on the task. The general formulas for computing Cohen's *d* and its variance are shown in (1) and (2), respectively. In (1) the mean difference between groups is divided by the pooled standard deviation. The correction factor '*J*' (Borenstein et al., 2011) is shown in (3), and finally the equations for Cohen's *d* and its variance that take the correction factor into account are shown in (4) and (5), respectively.

$$d = \frac{\bar{x}_{control} - \bar{x}_{study}}{SD_{pooled\ within}} \tag{1}$$

$$Var(d) = \frac{n_{control} + n_{study}}{n_{control} \times n_{study}} + \frac{d^2}{2(n_{control} + n_{study})} \tag{2}$$

$$J = 1 - \frac{3}{4df - 1} \tag{3}$$

$$d = J \times d \tag{4}$$

$$Var(d) = J^2 \times Var(d) \tag{5}$$

Where:

$$df = n_{control} + n_{study} - 2$$

$\bar{x}$ = Mean difference in RTs between the final random block and the preceding sequence block.

$SD_{\text{pooled within}}$ = Within-group *SD* of the difference between the final random block and preceding sequence block, pooled across the control and study groups.

The result from each study included in the meta-analysis was described using a single effect size that quantified the comparison between the groups on the difference in RT between the random block and the preceding sequence block. For nine studies, this was obtained from the reported *F*-ratio that tested for the Group × Block interaction (Cameli, 2006; Ferraro et al., 1993; Gawrys et al., 2008; Gilbert, 2003; Muslimovic et al., 2007; Sarazin, Deweer, Pillon, Merkl, & Dubois, 2001; Smith & McDowall, 2004; Smith et al., 2001; Westwater et al., 1998). In one case, the reported value from an independent measures *t*-test was the data extracted (van Tilborg & Hulstijn, 2010). The studies by Jackson, Jackson, Harrison, Henderson, and Kennard (1995), Selco (1998), and Sommer, Grafman, Clark, and Hallett (1999) reported *M* and *SD* of the difference between the random and preceding sequence block for each group. For a further six studies, data was extracted from

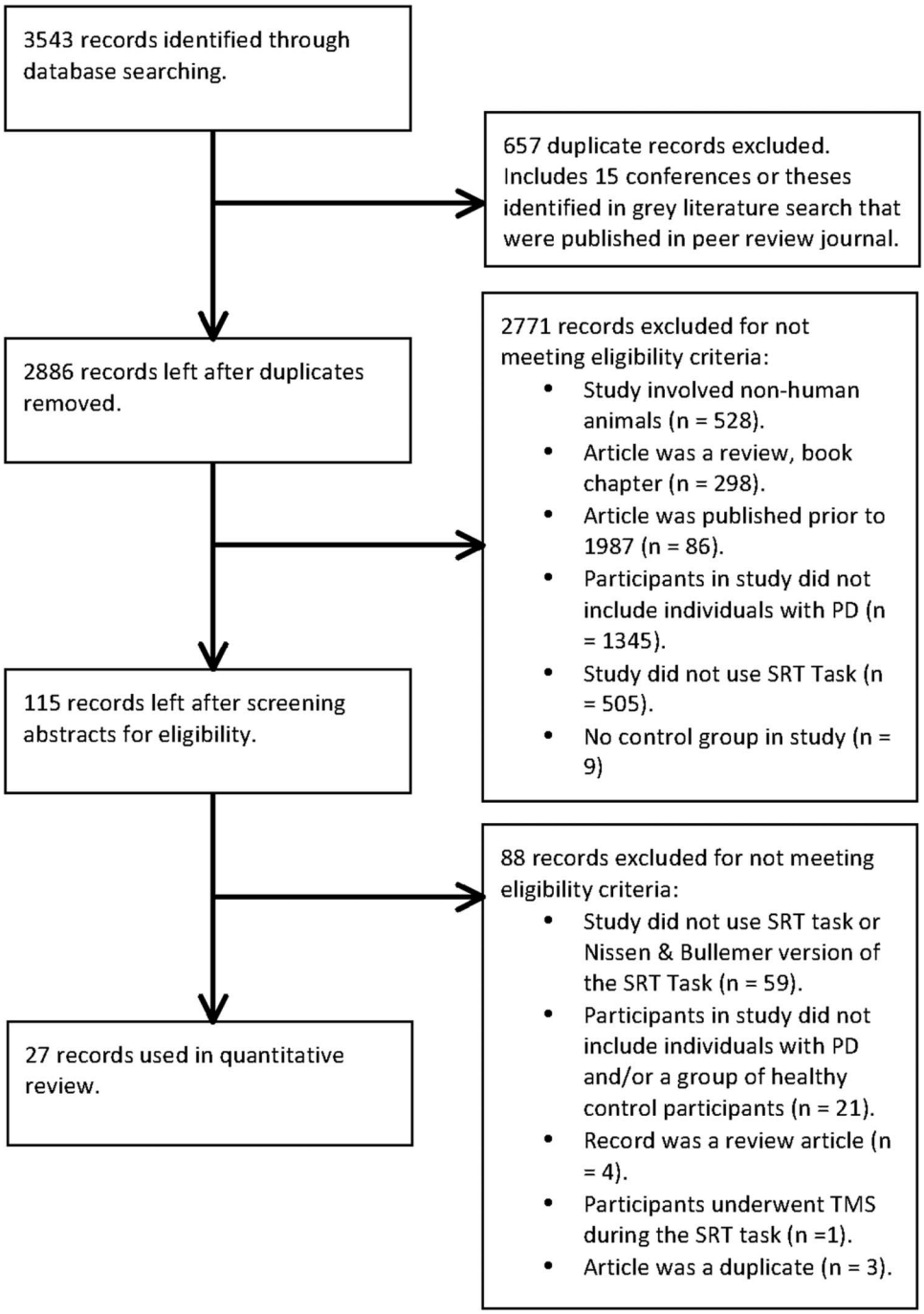


Figure 1. PRISMA Flowchart showing process of identifying articles included in the meta-analysis.

figures (Brown et al., 2003; Deroost et al., 2006; Smith & McDowall, 2006) or a combination of a figure and reported values in text (Kelly et al., 2004; Shin & Ivry, 2003; Vandenbossche et al., 2013). Means were extracted from figures and the standard deviations calculated using a reported *t*-statistic or *F*-statistic for the studies by Helmuth, Mayr, and Daum (2000), Werheid, Ziessler, Nattkemper, and Von Cramon (2003), and Werheid, Zysset, Muller, Reuter, and Von Cramon (2003). Data presented in figures were converted using Plot Digitizer Software (Version 2.6.4). Finally, for five studies, data for *M* and *SD* of both sequence and random block were extracted from either a figure (Pascual-Leone et al., 1993; Seidler et al., 2007; Wang et al., 2009) or text (Bondi, 1991; Stefanova et al., 2000), and an estimate of the correlation between these blocks was used to calculate the effect size.

For five of the 27 studies included in the meta-analysis, it was necessary to combine two sets of effect sizes. In the studies by

Table 1
Summary of Study's Participant Characteristics

Study	Sample size		Mean age (years)		Mean PD symptom severity (Hoehn-Yahr scale)
	Study (n_{study})	Control ($n_{control}$)	Study group	Control group	
Bondi (1991)	19	19	67.3	69.3	2
Brown et al. (2003)	10	10	54.9	57.2	Not reported
Cameli (2006)	9	9	65	66	1.83[b]
Deroost et al. (2006)	16	16	66.6	65.9	3
Ferraro et al. (1993)	17	26	69	70	Not reported
Gawrys et al. (2008)	16	20	57	55.7	1.9
Gilbert (2003)	10	10	65.2	64.5	1.9
Helmuth, Mayr, & Daum (2000)	21	24	58.8	64.6	Not reported
Jackson et al. (1995)	11	10	67	67.5	Not reported
Kelly et al. (2004)	12	9	64.2	65.2	1.83
Muslimovic et al. (2007)	95	44	64.9	60.7	1.91
Pascual-Leone et al. (1993)	20	30	56	57	Not reported
Sarazin et al. (2001)	20	15	61.9	61.9	2.8
Seidler et al. (2007)	8	6	57.4	59.2	Not reported
Selco (1998)	12	10	65.6	67.3	2.09
Shin & Ivry (2003)	10	10	64	71	Not reported
Smith & McDowall (2004)	19	31	66.2	68	2.63
Smith & McDowall (2006)	29	28	63	65.7	2.32
Smith et al. (2001)	13	14	66.4	68.4	2.31
Sommer et al. (1999)	11	15	55.9	51.7	Not reported
Stefanova et al. (2000)	39	31	49.3	48.3	1.6
van Tilborg & Hulstijn (2010)	9	12	67.5	69.6	1.58
Vandenbossche et al. (2013)	28[a]	14	66.9	67.1	2.43
Wang, Sun, & Ding (2009)	20	20	57.6	Not reported	1.8
Werheid, Ziessler, et al. (2003)	11	11	60.3	59.3	2.45
Werheid, Zysset, et al. (2003)	7	7	58.7	52.9	1.5
Westwater et al. (1998)	13	9	62	60	1.67

[a] Comprises two PD subgroups. [b] Study reported that seven participants were in Stage 1–2 on the Hoehn-Yahr scale, and so in calculating the average it was taken as seven participants in Stage 1.5.

Deroost et al. (2006), Kelly, Jahanshahi, and Dirnberger (2004), and Smith and McDowall (2006), effect sizes were averaged from separate analyses that compared the PD and control group on two different types of sequence. In the Selco (1998) study, effect sizes were averaged from separate analyses that compared the PD and control group under conditions requiring a verbal response or a button-press response. In the Vandenbossche et al. (2013) study, effect sizes were averaged from separate analyses that compared two PD subgroups and the control group under single task and dual task conditions.

For all studies included in the meta-analysis ($n = 27$), data and moderator variables (presented in Tables 1 and 2) were independently extracted by both reviewers. This process was undertaken to check the reliability of data extracted from papers included in the meta-analysis. For all categorical and continuous moderator data presented in Tables 1 and 2 the reviewers were found to extract the same information. For data extracted from figures, reliability was checked by computing the correlation between reviewers' values. This value was found to be high ($r = .98$).

Comprehensive Meta-Analysis Software (Borenstein, Hedges, Higgins, & Rothstein, 2005) was used to convert the extracted data to a common effect size and variance. Description of the data extracted from each study and the method used in Comprehensive Meta-Analysis Software to convert extracted data to Cohen's *d* and *Var*(*d*) are described in the online supplemental materials.

Meta-Analytic Procedures

Two approaches were used in this report to synthesize the SRT task literature investigating procedural learning in PD. First, meta-analysis was used to compute an average effect size that quantified the overall difference between PD and Controls on the SRT task. An alpha level of 0.05 (two-tailed) was used to evaluate whether the average effect size was significantly different from zero. Effect sizes were averaged using a random effects model (Hedges & Olkin, 1985). This method assumes that differences or heterogeneity between study effect sizes is the sum of sampling error (referred to as within-study error) and systematic influences (referred to as between-study error or true heterogeneity). The percentage of heterogeneity attributable to between-study error was measured using the I^2 statistic. This statistic expresses, as a percentage, heterogeneity between study effect sizes due to between-study error. Larger I^2 values indicate the presence of systematic influences on study findings. As a general guideline it has been suggested that values of 25%, 50%, and 75% correspond to low, medium, and high levels of heterogeneity, respectively (Higgins, Thompson, Deeks, & Altman, 2003).

The second set of analyses used in this report investigated predictor variables that might account for between-study error or systematic influences that are related to differences between study level effect sizes. Predictor variables used in these analyses were study level characteristics presented in Tables 1 and 2. These data

Table 2
Summary of Study's SRT Task Characteristics

Study	Sequence	Proportion of ambiguous transitions[c]	Sequence length	Exposures to sequence[f]	Single or dual task	Response method
Bondi (1991)	4–2-3–1-3–2-4–3-2–1	0.30	10	16	Single	Key press
Brown et al. (2003)	4–3-1–2-4–1-3–1-4–2	0.60	10	48	Single	Key press
Cameli (2006)	4–2-3–1-3–2-4–3-2–1	0.30	10	40	Single	Verbal
Deroost et al. (2006)	1–2-1–3-4–2-3–1-4–3-2–4 & 1–3-2–3-4–2-1–3-4–1-4–2[a]	0.50[d]	12	56.25	Single	Key press
Ferraro et al. (1993)	4–2-3–1-3–2-4–3-2–1	0.30	10	40	Single	Key press
Gawrys et al. (2008)	1–2-1–4-2–3-4–1-3–2-4–3	1.00	12	20	Single	Key press
Gilbert (2004)	4–2-3–1-3–2-4–3-2–1	0.30	10	40	Single	Key press
Helmuth, Mayr, & Daum (2000)	1–3-4–2-3–2 & 1–2-4–3-4–2-3[b]	0.00[d]	6.5[e]	123.5[g]	Single	Key press
Jackson et al. (1995)	1–2-4–3-1–4-2–1-3–4-3	0.27	11	36	Single	Key press
Kelly et al. (2004)	4–2-3–4-2–1-3– 2-4–1 & 4–1-3–2-1–4-2–4-3–2[a]	0.30[d]	10	50	Dual	Key press
Muslimovic et al. (2007)	1–2-4–3-4–2-1–4-1–3	0.60	10	50	Single	Key press
Pascual-Leone et al. (1993)	4–2-3–1-3–2-4–3-2–1	0.30	10	40	Single	Key press
Sarazin et al. (2001)	Not reported		10	60	Single	Key press
Seidler et al. (2007)	Not reported		12	31.33	Dual	Key press
Selco (1998)	2–1-4–3-4–1-2–3-1–3-2–4 & 3–1-2–4-2–1-3–4-1–4-3–2[a]	1.00[d]	12	66	Single	Both
Shin & Ivry (2003)	1–4-2–1-3–2-4–3	0.00	8	108.5[g]	Single	Key press
Smith & McDowall (2004)	1–2-1–4-2–3-4–1-3–2-4–3 & 1–4-1–3-4–2-3–2-1–3-4–2[a]	0.50[d]	12	40	Single	Verbal
Smith & McDowall (2006)	1–4-2–1-3–2-4–3 & 4–2-4–1-2–3-1–3[b]	0.00[d]	8	66[g]	Single	Verbal
Smith et al. (2001)	4–2-3–1-3–2-4–3-2–1	0.30	10	40	Single	Verbal
Sommer et al. (1999)	3–2-4–1-2–3-4–3-2–1	0.30	10	40	Single	Key press
Stefanova et al. (2000)	1–3-1–2-4–3-2–4-1–3	0.00	10	40	Single	Key press
van Tilborg & Hulstijn (2010)	4–2-3–1-3–2-4–3-2–1	0.30	10	40	Single	Key press
Vandenbossche et al. (2013)	1–3-2–3-4–2-1–3-4–1-4–2	0.00	12	60	Both	Key press
Wang, Sun, & Ding (2009)	Not reported		Not reported		Single	Key press
Werheid, Ziessler, et al. (2003)	1–3-4–2-1–2-4–3	0.00	8	24	Single	Key press
Werheid, Zysset, et al. (2003)	1–2-1–4-2–3-4–1-3–2-4–3	1.00	12	120	Single	Key press
Westwater et al. (1998)	4–2-3–1-3–2-4–3-2–1	0.30	10	40	Single	Verbal

[a] Participants were presented with both sequences. [b] Participants were trained on one of two sequences. [c] Calculated as the number of ambiguous transitions divided into the sequence length. [d] Value averages proportion of ambiguous transitions across two sequences. [e] Value averages two sequences of different lengths. [f] Number of times sequence is presented prior to Random Block. [g] Value averages two different numbers of exposures.

were analyzed using random effects meta-regression (Greenland, 1987; Thompson & Higgins, 2002). Meta-regression tests whether one or more predictor variables significantly predict study level effect sizes.

Results

Evaluation of Publication Bias

Preliminary analyses investigated publication bias in the studies identified by the search criteria. Evidence of publication bias was evaluated using a funnel plot which is presented in Figure 2.

Funnel plots show the relationship between individual study effect sizes and sample size (or in some cases standard error). The sample size is taken as a measure of the precision or accuracy of a study's effect size. Studies with smaller samples have poorer precision. A funnel plot indicates the presence of publication bias if study effect sizes with low precision are asymmetrically distributed around the weighted average effect size (see Egger, Smith, Schneider, & Minder, 1997). Figure 2 shows that overall, effect sizes are symmetrically distributed around the weighted average effect size. The presence of publication bias was formally tested using Egger's Test, which indicated that effect sizes were not significantly asymmetrically distributed, intercept = 0.985, $t(25) = 0.855$, $p = .401$.

The presence of publication bias was also examined by testing whether there was a significant difference in effect sizes between published and unpublished studies. For published studies the average effect size was found to be .554 and for unpublished studies .507. The difference between these effect sizes was not significant ($p = .890$). As a further test of publication bias, a "classic fail-safe N" value was computed (also known as "file-drawer analysis"). This value indicates the number of nonsignificant studies, not included in this meta-analysis, required to bring alpha to .05. This analysis revealed that a total of 384 nonsignificant studies are required. Thus, publication bias seems unlikely.

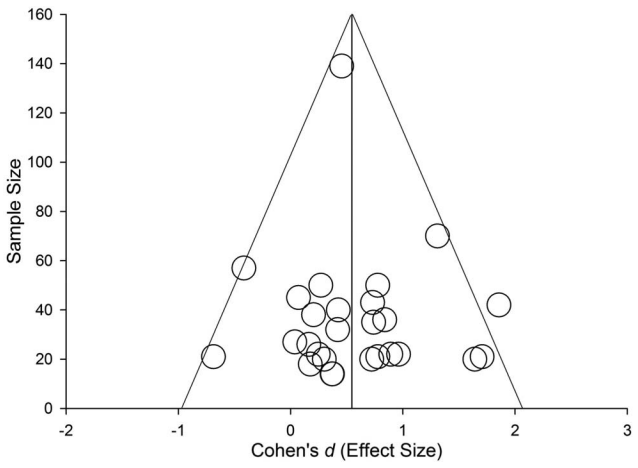


Figure 2. Funnel plot showing study level effect sizes plotted against sample size. Evidence of publication bias evident when effect sizes are asymmetrically distributed around average effect size when study precision is low (i.e., sample size is small).

SRT Task Performance in Parkinson's Disease

The next analysis summarized results of individual studies to test whether, on average, there was a significant difference between PD and controls on sequence learning in the SRT task. A forest plot showing individual study effect sizes, as well as the weighted average effect size (along with 95% confidence intervals for the study and average weighted effect sizes), is presented in Figure 3. As noted earlier, effect sizes were averaged using a random effects model (Hedges & Olkin, 1985). Positive *d* values indicate that the control group had a larger difference in RTs between sequenced and random blocks as compared with the PD group. That is, positive values indicate that the PD group performed poorly on the SRT task, as compared with the control group. The weighted average effect size was found to be .531 and was statistically significant ($p < .001$; unweighted average = .570; unweighted median = .426). The magnitude of the average effect size indicates that the difference in RTs between random and sequence blocks on the SRT task is about half a standard deviation larger in control groups compared with PD groups. According to Cohen's (1988) taxonomy this value corresponds to a medium effect size.

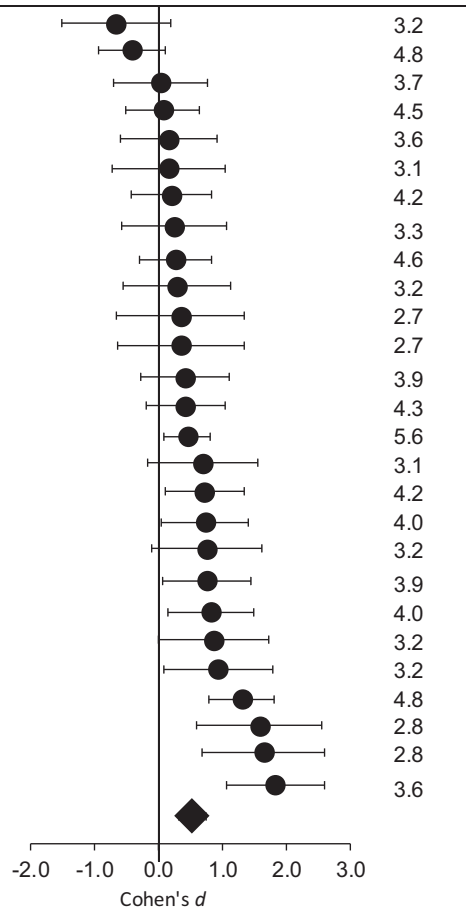
Even though the average effect size is significant, there is considerable variability between individual study effect sizes. Effect sizes ranged from 1.822 (Vandenbossche et al., 2013) to −0.659 (Kelly et al., 2004); negative values indicate that the control group performed worse on the SRT task than the PD group. Calculation of the I^2 statistic for the studies in Figure 3 was found to be 58. This indicates that 58% of variability between effect sizes reflects the influence of between-study error or systematic influence on the data. Alternatively stated, this result indicates that 42% of differences in effect sizes can be attributable to within-study error or chance. The I^2 value of 58 indicates medium to high levels of heterogeneity between study findings (Higgins et al., 2003).

Investigating the Sources of Heterogeneity in Study Findings

The next set of analyses used random effects meta-regression to investigate the sources of the between-study error/systematic influence (i.e., the percentage of heterogeneity between effect sizes measured by the I^2 statistic). Specifically, we tested whether participant and SRT task methodological characteristics, presented in Tables 1 and 2, predicted the study effect sizes that are presented in Figure 3.

There were an insufficient number of studies to test all covariates in a single model. In meta-regression, a ratio of 10 studies for each predictor variable is recommended (Borenstein et al., 2011). Therefore, separate meta-regressions were performed to investigate one predictor variable at time. Interactions between predictor variables were also investigated, again in separate analyses. The interaction term was created by multiplying constituent variables. Continuous variables were centered prior to multiplication. Because there were an insufficient number of studies in the meta-analysis to simultaneously test main and interaction effects, interaction terms were tested by first removing the influence of the main effects, before being entered into the model. This was achieved using least squares regression. Specifically, each interaction term was regressed onto its constituent variables and stan-

Study	Cohen's *d*	Variance	95% C.I. Lower	95% C.I. Upper	*p*-value	Weight (%)
Kelly et al. (2004)	-0.659[a]	0.192	-1.517	0.200	.132	3.2
Smith & McDowall (2006)	-0.410	0.070	-0.928	0.108	.121	4.8
Smith et al. (2001)	0.037	0.140	-0.695	0.769	.920	3.7
Helmuth, Mayr, & Daum (2000)	0.070	0.086	-0.505	0.646	.811	4.5
Sommer et al. (1999)	0.159	0.148	-0.596	0.913	.680	3.6
Cameli (2006)	0.168	0.202	-0.714	1.050	.709	3.1
Bondi (1991)	0.202	0.101	-0.423	0.826	.527	4.2
Selco (1998)	0.238[b]	0.172	-0.575	1.050	.566	3.3
Pascual-Leone et al. (1993)	0.266	0.081	-0.293	0.826	.351	4.6
Brown et al. (2003)	0.288	0.186	-0.556	1.133	.503	3.2
Seidler et al. (2007)	0.347	0.260	-0.653	1.346	.497	2.7
Werheid, Zysset, et al. (2003)	0.350	0.255	-0.639	1.340	.488	2.7
Deroost et al. (2006)	0.410[c]	0.124	-0.281	1.100	.245	3.9
Wang, Sun, & Ding (2009)	0.418	0.098	-0.197	1.032	.183	4.3
Muslimovic et al. (2007)	0.454	0.034	0.095	0.814	.013*	5.6
Shin & Ivry (2003)	0.692	0.195	-0.174	1.559	.117	3.1
Ferraro et al. (1993)	0.718	0.100	0.099	1.337	.023*	4.2
Sarazin et al. (2001)	0.724	0.119	0.048	1.399	.036*	4.0
van Tilborg & Hulstijn (2010)	0.747	0.192	-0.113	1.607	.089	3.2
Smith & McDowall (2004)	0.758[c]	0.125	0.065	1.451	.032*	3.9
Gawrys et al. (2008)	0.821	0.117	0.151	1.491	.016*	4.0
Westwater et al. (1998)	0.860	0.191	0.004	1.716	.049*	3.2
Werheid, Ziessler, et al. (2003)	0.926	0.188	0.076	1.775	.033*	3.2
Stefanova et al. (2000)	1.295	0.069	0.781	1.808	<.001**	4.8
Gilbert (2004)	1.573	0.245	0.602	2.544	.002*	2.8
Jackson et al. (1995)	1.639	0.240	0.679	2.599	<.001**	2.8
Vandenbossche et al. (2013)	1.822[d]	0.150	1.063	2.581	<.001**	3.6
Average	**0.531**	**0.011**	**0.322**	**0.740**	**<.001****	



Notes: [a]Effect size averages results from groups' performance on ambiguous and hybrid sequences; [b]Effect size averages results from groups' performance on SRT task using verbal and keypress response methods [c]Effect size averages results from groups' performance on FOC and SOC sequences; [d]Effect size average groups' performance on SRT task completed under single-task conditions and dual-task conditions.
*$p < .05$; **$p < .001$

Figure 3. Forest plot showing study level and average weighted effect sizes.

dardized residuals were saved. The standardized residuals were then entered into the model. These residuals represent the interaction term after removing covariance related to each main effect.

The predictor variables tested were the average severity of PD symptoms (from Table 1), the method used to collect responses on the task (response box input vs. voice input), whether the SRT task was presented under single task or dual task conditions, the type of sequence used (measured by the proportion of ambiguous transitions in the sequence; see Method), the length of the sequence, and the number of times it was presented (all from Table 2). Predictor variables representing binary variables (e.g., response method, testing condition) were dummy coded. Specifically, response method was coded so that 0 = verbal response and 1 = keyboard/response box, and testing condition was coded so that 0 = single task and 1 = dual task.

The outcome variable in the meta-regression analyses were study effect sizes that are presented in Figure 3. However, for the meta-regression analyses testing whether single versus dual task procedures predicted effect sizes it was necessary to choose one of two effect sizes from Vandenbossche et al. (2013). Similarly, it was necessary to choose one of two effect sizes from Selco (1998) for the meta-regression analyses testing whether verbal versus keypress response method conditions predicted effect sizes. In those studies separate effect sizes were available for performance under single and dual task conditions, or for verbal or motor response conditions. In the meta-regression analyses only the effect size from the dual task condition was used from the Vandenbossche et al. (2013) study, and only the effect size from the verbal response method condition was used from Selco (1998). These effect sizes were selected to increase the number of data points available for the dual task conditions and the verbal response conditions. It was not possible to use both effect sizes from each of these studies because this would require treating dependent data points as nondependent. The results from the meta-regression analyses are presented in Table 3.

Table 3
Summary of Coefficients From Meta-Regression

Model no.	Variable	*k*	*df*	Q_{Model}	$Q_{Residual}$	R^2	β	*p*
1	Single or dual task	27	1,25	0.437	27.771	0.016	0.124	.509
2	Proportion of ambiguous transitions	24	1,22	0.288	23.402	0.012	−0.110	.592
3	No. of exposures to sequence	26	1,24	1.060	24.279	0.042	−0.205	.303
4	Response method	27	1,25	2.500	24.072	0.094	0.307	.114
5	Sequence length	26	1,24	2.467	22.866	0.097	0.312	.116
6	Symptom severity	19	1,17	0.002	18.278	0.000	−0.009	.969
7	Single vs. Dual Task × Prop. of Ambiguous Transitions	24	1,22	11.906	13.424	0.470	−0.686	<.001**
8	Single vs. Dual Task × No. of Exposures to Sequence	26	1,24	4.671	22.351	0.173	0.416	.031*
9	Single vs. Dual Task × Response Method	27	1,25	1.157	25.283	0.044	−0.209	.282
10	Single vs. Dual Task × Sequence Length	26	1,24	5.608	21.414	0.208	0.456	.018*
11	Single vs. Dual Task × Symptom Severity	19	1,17	11.213	8.693	0.563	0.751	<.001**
12	Proportion of Ambiguous Transitions × No. of Exposures to Sequence	24	1,22	0.403	23.287	0.017	−0.130	.526
13	Proportion of Ambiguous Transitions × Response Method	24	1,22	0.495	23.320	0.021	−0.144	.482
14	Proportion of Ambiguous Transitions × Sequence Length	24	1,22	2.009	21.681	0.085	−0.291	.156
15	Proportion of Ambiguous Transitions × Symptom Severity	17	1,15	0.197	16.004	0.012	0.110	.657
16	No. of Exposures to Sequence × Response Method	26	1,24	1.774	24.754	0.067	0.259	.183
17	No. of Exposures to Sequence × Sequence Length	26	1,24	0.674	24.665	0.027	0.163	.412
18	No. of Exposures to Sequence × Symptom Severity	18	1,16	0.191	16.974	0.011	0.105	.662
19	Response Method × Sequence Length	26	1,24	0.497	26.030	0.019	−0.137	.481
20	Response Method × Symptom Severity	19	1,17	0.267	18.133	0.015	0.120	.606
21	Sequence Length × Symptom Severity	18	1,16	0.713	16.452	0.042	0.204	.399

* $p < .05$. ** $p < .001$.

None of the main effects were found to be significant predictors of effect sizes (Models 1–6). The analyses testing interaction terms were also all nonsignificant, except for models that included the single versus dual task predictor variable in the interaction term. Models 7, 8, 10, and 11, tested the interaction between single versus dual task and proportion of ambiguous transitions (Model 7), number of exposures to the sequence (Model 8), sequence length (Model 10), and PD symptom severity (Model 11). For Models 8, 10, and 11 the beta-value was found to be positive. This indicates that larger positive effect size values (i.e., larger differences between groups) are observed for studies that use SRT tasks that provided more exposures to the target sequence (Model 8), have longer sequences (Model 10), or have participants with more severe PD symptoms (Model 11), but only under dual task conditions. The model testing the dual versus Single Task × Proportion of Ambiguous Transitions (Model 7) was also significant but the beta-value was negative. This result indicates effect sizes become smaller (i.e., the difference between PD and control groups decreases) for studies that present sequences with more ambiguous transitions, but only under dual task conditions. For illustrative purposes, significant models (Models 7, 8, 10, and 11) are presented in Figure 4.

Discussion

In this report, a meta-analysis was undertaken to quantitatively summarize research investigating the performance of individuals with PD and neurologically intact controls on implicit sequence learning in SRT tasks. The results from 27 studies, representing data from 505 participants with PD and 460 controls, were included in the meta-analysis. The weighted average effect size, computed using a random effects model, was 0.531 (95% CI [.322, .740]), and was statistically significant. This is a medium effect size according to Cohen's (1988) guidelines. This provides strong evidence that sequence learning in procedural memory is impaired or dysfunctional in PD. The results of the meta-analysis provide further evidence supporting the view that implicit learning on the SRT task is sensitive to basal ganglia pathology (Kandel et al., 2012; Packard & Knowlton, 2002; Parent & Hazrati, 1995; Ullman, 2004).

The average effect size observed in this report is consistent with Siegert et al.'s (2006) result. When using a random effects model, they observed an average effect size of 0.65 (95% CI [0.10, 1.20]). According to Cohen's (1988) taxonomy this value also corresponds to a medium effect size. However, because Siegert et al. (2006) were only able to include six studies in their review, each with relatively small sample sizes, the width of the confidence interval for the average effect size was large; 1.1 *SD* units. By comparison, the width of the confidence interval in the current meta-analysis was .42 *SD* units. The current report builds on the previous meta-analysis by providing a more precise estimate of the average effect size.

Interestingly, meta-analyses investigating nonmotor areas of cognitive functioning in PD have also often reported medium average effect sizes. Siegert, Weatherall, Taylor, and Abernethy (2008) synthesized results of studies that investigated working memory in individuals with PD, as compared with neurologically intact controls. Medium average effect sizes, ranging from .56 to .74, were reported for measures of visual working memory, with a small effect size of .22 reported for measures of verbal working memory. Similarly, Kudlicka, Clare, and Hindle (2011) reported medium to large average effect sizes across varied measures of executive functioning. PD also appears to be associated with a deficit in recognizing the emotion of another person based on facial expression and tone of voice. In a meta-analysis, Gray and Tickle-Degnen (2010) reported that individuals with PD perform on average .52 standard deviations below their peers on measures of emotion recognition. In each of the meta-analyses mentioned, individuals with PD perform more poorly than their

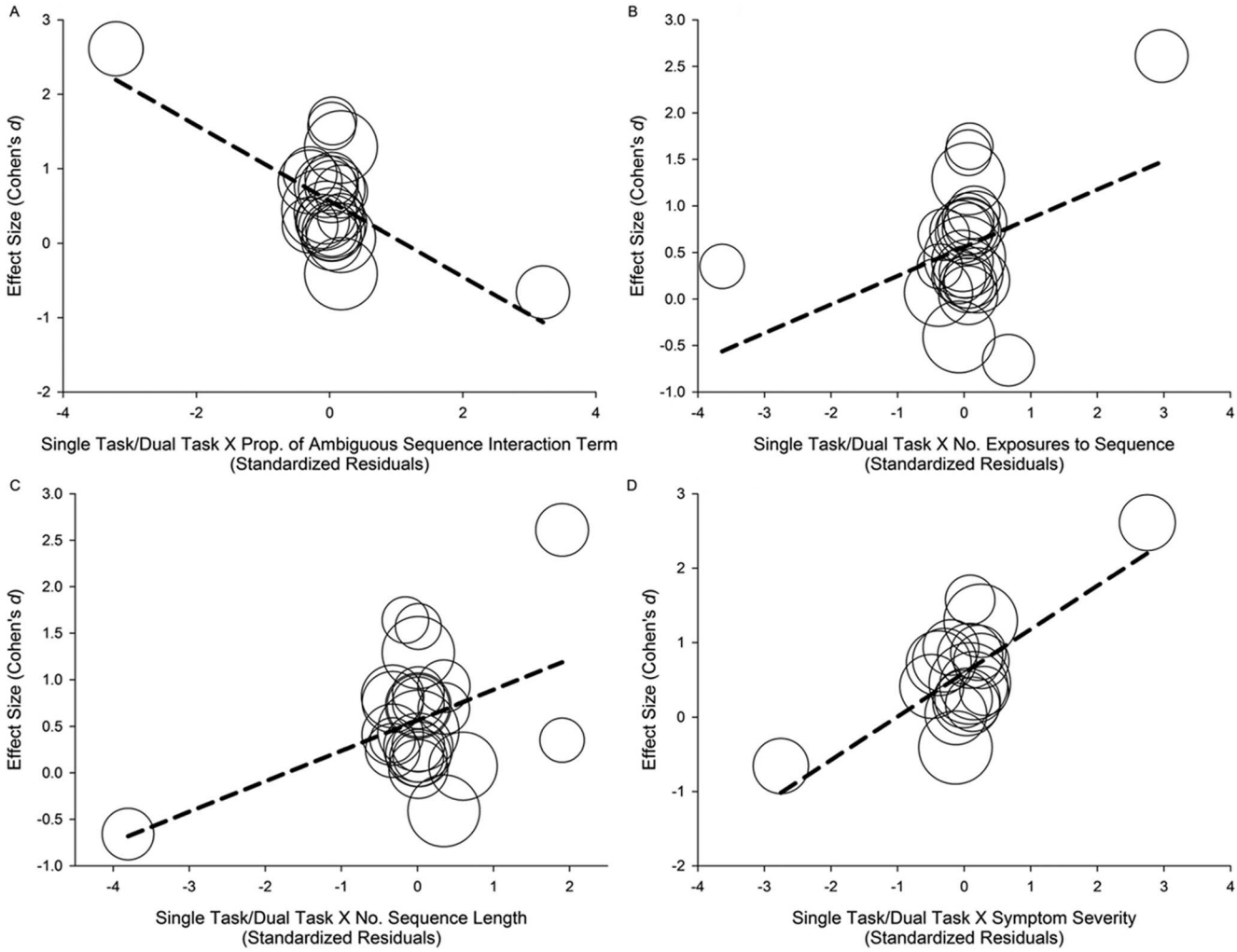


Figure 4. Scatterplot showing observed and predicted effect sizes. Panel (A) predictor variable is 'Single/Dual Task' X Proportion of Ambiguous Sequence' interaction term. Panel (B) predictor variable is 'Single/Dual Task' X Number of Exposures to Sequence. Panel (C) predictor variable is 'Single/Dual Task' X Sequence Length. Panel (D) predictor variable is 'Single/Dual Task' X Symptom Severity. Data points are proportionally sized to their weight in the model.

peers by a similar magnitude to that found in the current report. Thus, procedural learning is one of several cognitive functions affected in PD.

The analyses presented in this report add to the literature by testing which systematic influences might be related to differences in study findings. Both the current meta-analysis and that published by Siegert et al. (2006) revealed medium to high levels of between-study error/systematic influences on study findings. The value of the I^2 observed by Siegert et al. was 64.8, and in the current report, it was 58. We used meta-regression to investigate a number of variables that could explain the between-study differences. Few of the meta-regression models were found to be significant predictors of study effect sizes. The models that tested main effects of participant and methodological aspects of the SRT task on study effect sizes were not found to be significant (Table 3, Models 1–6). Models testing interactions between characteristics of the sequence, method for collecting responses, and severity of PD symptoms (Table 3, Models 12–21) were also not found to be significant predictors of study effect sizes.

However, four models that included the single/dual task conditions variable in the interaction term were found to be significant (Models 7, 8, 10, and 11). The results of three models (Models 8, 10, and 11) indicated that studies observed a larger effect size (i.e., bigger difference between PD and controls) when participants completed the SRT task under dual task conditions and when the sequence used on the SRT task was longer (Model 10), was presented more times relatively to other studies (Model 8), or when the PD participants had more severe symptoms (Model 11). One model (Model 7) indicated that studies were likely to observe smaller effect sizes (i.e., smaller difference between PD and control groups) when completing the SRT task under dual task conditions and when the sequence used comprised a higher proportion of ambiguous sequences.

In interpreting the significant meta-regression models it is important to note that the linear associations between predictor variable and effect sizes are largely influenced by the findings of two studies (see Figure 4). Furthermore in meta-regression there are other variables that may correlate with the predictor variables that may contribute to the significant result (Thompson & Higgins, 2002). Thus, explanations for the significant meta-regression models are offered tentatively.

The meta-regression models showing a larger difference between the PD and control groups when completing the SRT task under dual task conditions are consistent with some previous research. Several studies have shown that individuals with the disorder perform worse than neurologically intact controls when completing a motor or cognitive task under dual task conditions (Dalrymple-Alford et al., 1994; Wu & Hallett, 2008). In accounting for this pattern of results it has been suggested that PD is associated with cognitive deficits that limit the efficacy of switching between tasks or the amount of cognitive resources available to process information (Brown & Marsden, 1991). Thus, results from the meta-regression might indicate that the effect of dual task conditions on cognitive functioning may extend to the implicit learning of visuospatial information as well, possibly only in the presence of additional factors that increase task difficulty (longer sequences or more severe PD). However, this claim seems to be tempered by the results from Models 7 and 8, which showed an opposite trend. Model 8 showed larger differences between PD and control groups when the sequence was presented more often under dual task conditions. One explanation is that dual task conditions limit the extent to which PD patients can take advantage of the increased stimulus presentations, relative to controls, leading to larger effect sizes. Model 7 indicated smaller differences between PD and control groups when the SRT task was completed under dual task conditions and when the sequence used comprised a higher proportion of ambiguous transitions. An explanation for this result is that under this set of conditions, both the PD and control groups perform poorly, perhaps because dual task conditions can impede learning in the medial temporal lobe (Foerde, Knowlton, & Poldrack, 2006), which appears to be critical for learning ambiguous sequences (see beginning of article); poor performance by both groups in these circumstances could potentially lead to smaller effect sizes. However, equally plausible is that this association represents the influence of another variable correlated with studies investigating dual task performance in PD. Finally, as noted earlier, the significant meta-regression models appear to reflect the influence of two studies. Thus, we emphasize that these suggestions regarding the role of dual task paradigms on SRT task performance should be verified via experimental research.

Conclusion

This report used meta-analysis and meta-regression to examine the performance of individuals with PD on implicit sequence learning in the SRT task. The meta-analysis included 27 studies, representing data from 505 individuals with PD and 460 neurologically intact controls. It was found that individuals with PD perform just over half a standard deviation worse than controls on sequence learning in the task. Substantial variability was observed between study effect sizes; this variability appears to be related to whether the task was administered under single or dual task conditions. However, experimental work will now be required to test these suggestions.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Bondi, M. W. (1991). Neurobehavioural functioning in Parkinson's disease: The role of basal ganglia-thalamocortical circuit loops in predicting performance. (Doctor of Philosophy), The University of Arizona, Tucson, AZ.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). *Comprehensive meta-analysis (Version 2)*. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. Hoboken, NJ: Wiley.

*Brown, R. G., Jahanshahi, M., Limousin-Dowsey, P., Thomas, D., Quinn, N. P., & Rothwell, J. C. (2003). Pallidotomy and incidental sequence learning in Parkinson's disease. *NeuroReport, 14,* 21–24. doi:10.1097/00001756-200301200-00004

Brown, R. G., & Marsden, C. D. (1991). Dual task performance and processing resources in normal subjects and patients with Parkinson's disease. *Brain, 114,* 215–231.

*Cameli, L. (2006). Memory and language: Insights from picture description and past tense generation in a native and a second language in bilingual Alzheimer and Parkinson patients. (Doctor of Philosophy), Concordia University, Portland, Oregon.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17–30. doi:10.1037/0278-7393.16.1.17

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Laurence Erlbaum.

Cooper, J. A., Sagar, H. J., Jordan, N., Harvey, N. S., & Sullivan, E. V. (1991). Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. *Brain, 114,* 2095–2122. doi:10.1093/brain/114.5.2095

Curran, T. (1997). Higher-order associative learning in amnesia: Evidence from the serial reaction time task. *Journal of Cognitive Neuroscience, 9,* 522–533. doi:10.1162/jocn.1997.9.4.522

Dalrymple-Alford, J. C., Kalders, A. S., Jones, R. D., & Watson, R. W. (1994). A central executive deficit in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry, 57,* 360–367. doi:10.1136/jnnp.57.3.360

*Deroost, N., Kerckhofs, E., Coene, M., Wijnants, G., & Soetens, E. (2006). Learning sequence movements in a homogenous sample of patients with Parkinson's disease. *Neuropsychologia, 44,* 1653–1662. doi:10.1016/j.neuropsychologia.2006.03.021

Dubois, B., & Pillon, B. (1996). Cognitive deficits in Parkinson's disease. *Journal of Neurology, 244,* 2–8. doi:10.1007/PL00007725

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315,* 629–634. doi:10.1136/bmj.315.7109.629

Ergorul, C., & Eichenbaum, H. (2006). Essential role of the hippocampal formation in rapid learning of higher-order sequential associations. *The Journal of Neuroscience, 26,* 4111–4117. doi:10.1523/JNEUROSCI.0441-06.2006

*Ferraro, F. R., Balota, D. A., & Connor, L. T. (1993). Implicit memory and the formation of new associations in nondemented Parkinson's disease individuals and individuals with senile dementia of the Alzheimer type: A serial reaction time (SRT) investigation. *Brain and Cognition, 21,* 163–180. doi:10.1006/brcg.1993.1013

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

*Gawrys, L., Szatkowska, I., Jamrozik, Z., Janik, P., Friedman, A., & Kaczmarek, L. (2008). Nonverbal deficits in explicit and implicit memory of Parkinson's disease patients. *Acta Neurobiologiae Experimentalis, 68,* 58–72.

*Gilbert, B. (2003). Working memory and procedural memory in Parkinson's disease: Evaluation and training. (Doctor of Philosophy), University of Montreal, Montreal, Canada.

Gray, H. M., & Tickle-Degnen, L. (2010). A meta-analysis of performance on emotion recognition tasks in Parkinson's disease. *Neuropsychology, 24,* 176–191. doi:10.1037/a0018104

Greenland, S. (1987). Quantitative methods in the review of epidemiologic literature. *Epidemiologic Reviews, 9,* 1–30.

Hardwick, R. M., Rottschy, C., Miall, R. C., & Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *NeuroImage, 67,* 283–297. doi:10.1016/j.neuroimage.2012.11.020

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Helkala, E., Laulumaa, V., Soininen, H., & Riekkinen, P. J. (1988). Recall and recognition memory in patients with Alzheimer's and Parkinson's diseases. *Annals of Neurology, 24,* 214–217. doi:10.1002/ana.410240207

*Helmuth, L. L., Mayr, U., & Daum, I. (2000). Sequence learning in Parkinson's disease: A comparison of spatial-attention and number-response sequences. *Neuropsychologia, 38,* 1443–1451. doi:10.1016/S0028-3932(00)00059-2

Higgins, J., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557–560. doi:10.1136/bmj.327.7414.557

Hoehn, M. M., & Yahr, M. D. (1967). Parkinsonism: Onset, progression, and mortality. *Neurology, 17,* 427–442.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis*. Beverly Hills, CA: Sage Publications.

*Jackson, G. M., Jackson, S. R., Harrison, J., Henderson, L., & Kennard, C. (1995). Serial reaction time learning and Parkinson's disease: Evidence for a procedural learning deficit. *Neuropsychologia, 33,* 577–593. doi:10.1016/0028-3932(95)00010-Z

Jankovic, J. (2008). Parkinson's disease: Clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry, 79,* 368–376. doi:10.1136/jnnp.2007.131045

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2012). *Principles of neural science* (5th ed.). New York, NY: McGraw-Hill.

*Kelly, S. W., Jahanshahi, M., & Dirnberger, G. (2004). Learning of ambiguous versus hybrid sequences by patients with Parkinson's disease. *Neuropsychologia, 42,* 1350–1357. doi:10.1016/j.neuropsychologia.2004.02.013

Knopman, D., & Nissen, M. J. (1991). Procedural learning is impaired in Huntington's disease: Evidence from the serial reaction time task. *Neuropsychologia, 29,* 245–254. doi:10.1016/0028-3932(91)90085-M

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273,* 1399–1402. doi:10.1126/science.273.5280.1399

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

Kudlicka, A., Clare, L., & Hindle, J. V. (2011). Executive functions in Parkinson's disease: A systematic review and meta-analysis. *Movement Disorders, 26,* 2305–2315. doi:10.1002/mds.23868

Lang, A. E., & Lozano, A. M. (1998). Parkinson's disease. *The New England Journal of Medicine, 339,* 1044–1053. doi:10.1056/NEJM199810083391506

Lozano, A. M., Lang, A. E., Galvez-Jimenez, N., Miyasaki, J., Duff, J., Hutchison, W. D., & Dostrovsky, J. O. (1995). Effect of GPi pallidotomy on motor function in Parkinson's disease. *The Lancet, 346,* 1383–1387. doi:10.1016/S0140-6736(95)92404-3

Lum, J. A. G., & Kidd, E. (2012). An examination of the associations among multiple memory systems, past tense, and vocabulary in typically developing 5-year-old children. *Journal of Speech, Language and Hearing Research, 55,* 989–1006. doi:10.1044/1092-4388(2011/10-0137)

Lum, J. A. G., Ullman, M. T., & Conti-Ramsden, G. (2013). Procedural learning is impaired in dyslexia: Evidence from a meta-analysis of serial reaction time studies. *Research in Developmental Disabilities, 34,* 3460–3476. doi:10.1016/j.ridd.2013.07.017

Mayes, A., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences, 11,* 126–135. doi:10.1016/j.tics.2006.12.003

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*. doi:10.1371/journal.pmed.1000097

*Muslimovic, D., Post, B., Speelman, J. D., & Schmand, B. (2007). Motor procedural learning in Parkinson's disease. *Brain, 130,* 2887–2897. doi:10.1093/brain/awm211

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1–32. doi:10.1016/0010-0285(87)90002-8

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,* 519–532. doi:10.1016/S0028-3932(96)00101-7

Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience, 25,* 563–593. doi:10.1146/annurev.neuro.25.112701.142937

Parent, A., & Hazrati, L. (1995). Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Research Reviews, 20,* 91–127. doi:10.1016/0165-0173(94)00007-C

*Pascual-Leone, A., Grafman, J., Clark, K., Stewart, M., Massaquoi, S., Lou, J. S., & Hallett, M. (1993). Procedural learning in Parkinson's disease and cerebellar degeneration. *Annals of Neurology, 34,* 594–602. doi:10.1002/ana.410340414

Pascual-Leone, A., Valls-Sole, J., Brasil-Neto, J. P., Cammarota, A., Grafman, J., & Hallett, M. (1994). Akinesia in Parkinson's disease. II. Effects of subthreshold repetitive transcranial motor cortex stimulation. *Neurology, 44,* 892–898. doi:10.1212/WNL.44.5.892

Pascual-Leone, A., Wassermann, E. M., Grafman, J., & Hallett, M. (1996). The role of the dorsolateral prefrontal cortex in implicit procedural learning. *Experimental Brain Research, 107,* 479–485. doi:10.1007/BF00230427

Poldrack, R. A., & Rodriguez, P. (2003). Sequence learning: What's the hippocampus to do? *Neuron, 37,* 891–893. doi:10.1016/S0896-6273(03)00159-4

Price, A., & Shin, J. C. (2009). The impact of Parkinson's disease on sequence learning: Perceptual pattern learning and executive function. *Brain & Cognition, 69,* 252–261. doi:10.1016/j.bandc.2008.07.013

Robertson, E. M. (2007). The serial reaction time task: Implicit motor skill learning? *The Journal of Neuroscience, 27,* 10073–10075. doi:10.1523/JNEUROSCI.2747-07.2007

*Sarazin, M., Deweer, B., Pillon, B., Merkl, A., & Dubois, B. (2001). Procedural learning and Parkinson's disease: Implications of the striato-frontal loops. *Revue Neurologique, 157,* 1513–1518.

Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003). An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron, 37,* 1013–1025. doi:10.1016/S0896-6273(03)00123-5

*Seidler, R. D., Tuite, P., & Ashe, J. (2007). Selective impairments in implicit learning in Parkinson's disease. *Brain Research, 1137,* 104–110. doi:10.1016/j.brainres.2006.12.057

*Selco, S. L. (1998). *The nature of the processes, representations, and neural substrates which support and contribute to motor sequence*

learning. (Doctor of Philosophy), University of Illinois, Champaign, IL.

*Shin, J. C., & Ivry, R. B. (2003). Spatial and temporal sequence learning in patients with Parkinson's disease or cerebellar lesions. *Journal of Cognitive Neuroscience, 15*(8), 1232–1243. doi:10.1162/089892903322598175

Siegert, R., Taylor, K. D., Weatherall, M., & Abernethy, D. (2006). Is implicit sequence learning impaired in Parkinson's disease? A meta-analysis. *Neuropsychology, 20,* 490–495. doi:10.1037/0894-4105.20.4.490

Siegert, R., Weatherall, M., & Bell, E. M. (2008). Is implicit sequence learning impaired in schizophrenia? A meta-analysis. *Brain and Cognition, 67,* 351–359. doi:10.1016/j.bandc.2008.02.005

Siegert, R., Weatherall, M., Taylor, K. D., & Abernethy, D. (2008). A meta-analysis of performance on simple span and more complex working memory tasks in Parkinson's disease. *Neuropsychology, 22,* 450–461. doi:10.1037/0894-4105.22.4.450

*Smith, J. G., & McDowall, J. (2004). Impaired higher order implicit sequence learning on the verbal version of the serial reaction time task in patients with Parkinson's disease. *Neuropsychology, 18,* 679–691. doi:10.1037/0894-4105.18.4.679

*Smith, J. G., & McDowall, J. (2006). The implicit sequence learning deficit in patients with Parkinson's disease: A matter of impaired sequence integration? *Neuropsychologia, 44,* 275–288. doi:10.1016/j.neuropsychologia.2005.05.001

*Smith, J. G., Siegert, R., McDowall, J., & Abernethy, D. (2001). Preserved implicit learning on both the serial reaction time task and artificial grammar in patients with Parkinson's disease. *Brain and Cognition, 45,* 378–391. doi:10.1006/brcg.2001.1286

*Sommer, M., Grafman, J., Clark, K., & Hallett, M. (1999). Learning in Parkinson's disease: Eyeblink conditioning, declarative learning, and procedural learning. *Journal of Neurology, Neurosurgery and Psychiatry, 67,* 27–34. doi:10.1136/jnnp.67.1.27

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82,* 171–177. doi:10.1016/j.nlm.2004.06.005

Squire, L. R., Knowlton, B. J., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology, 44,* 453–495. doi:10.1146/annurev.ps.44.020193.002321

Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience, 27,* 279–306. doi:10.1146/annurev.neuro.27.070203.144130

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the United States of America, 93,* 13515–13522. doi:10.1073/pnas.93.24.13515

*Stefanova, E. D., Kostic, V. S., Ziropadja, L., Markovic, M., & Ocic, G. G. (2000). Visuomotor skill learning on serial reaction time task in patients with early Parkinson's disease. *Movement Disorders, 15,* 1095–1103. doi:10.1002/1531-8257(200011)15:6<1095::AID-MDS1006>3.0.CO;2-R

Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y., & Worden, M. S. (2004). Evidence of developmental differences in implicit sequence learning: An fMRI study of children and adults. *Journal of Cognitive Neuroscience, 16,* 1339–1351. doi:10.1162/0898929042304688

Thomas, K. M., & Nelson, C. A. (2001). Serial reaction time learning in preschool-and school-age children. *Journal of Experimental Child Psychology, 79,* 364–387. doi:10.1006/jecp.2000.2613

Thompson, S. G., & Higgins, J. P. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine, 21,* 1559–1573. doi:10.1002/sim.1187

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience, 2,* 717–726. doi:10.1038/35094573

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92,* 231–270. doi:10.1016/j.cognition.2003.10.008

Ullman, M. T. (2006). Is Broca's area part of a basal ganglia thalamocortical circuit? *Cortex, 42,* 480–485. doi:10.1016/S0010-9452(08)70382-4

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., & Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience, 9,* 266–276. doi:10.1162/jocn.1997.9.2.266

*Vandenbossche, J., Deroost, N., Soetens, E., Coomans, D., Spildooren, J., Vercruysse, S., . . . Kerckhofs, E. (2013). Impaired implicit sequence learning in Parkinson's disease patients with freezing of gait. *Neuropsychology, 27,* 28–36. doi:10.1037/a0031278

*van Tilborg, I., & Hulstijn, W. (2010). Implicit motor learning in patients with Parkinson's and Alzheimer's disease: Differences in learning abilities? *Motor Control, 14,* 344–361.

*Wang, X. P., Sun, B. M., & Ding, H. L. (2009). Changes of procedural learning in Chinese patients with non-demented Parkinson disease. *Neuroscience Letters, 449,* 161–163. doi:10.1016/j.neulet.2008.10.086

*Werheid, K., Ziessler, M., Nattkemper, D., & Von Cramon, D. Y. (2003). Sequence learning in Parkinson's disease: The effect of spatial stimulus-response compatibility. *Brain and Cognition, 52,* 239–249. doi:10.1016/S0278-2626(03)00076-9

*Werheid, K., Zysset, S., Muller, A., Reuter, M., & Von Cramon, D. Y. (2003). Rule learning in a serial reaction time task: An fMRI study on patients with early Parkinson's disease. *Cognitive Brain Research, 16,* 273–284. doi:10.1016/S0926-6410(02)00283-5

*Westwater, H., McDowall, J., Siegert, R., Mossman, S., & Abernethy, D. (1998). Implicit learning in Parkinson's disease: Evidence from a verbal version of the serial reaction time task. *Journal of Clinical and Experimental Neuropsychology, 20,* 413–418. doi:10.1076/jcen.20.3.413.826

Wu, T., & Hallett, M. (2008). Neural correlates of dual task performance in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 79,* 760–766. doi:10.1136/jnnp.2007.126599

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience, 7,* 464–476. doi:10.1038/nrn1919

Received November 18, 2013
Revision received May 27, 2014
Accepted May 27, 2014 ■