Theory of Mind Deficits in Parkinson's Disease: A Product of Executive Dysfunction?

Clare M. Eddy
Birmingham and Solihull Mental Health NHS Foundation Trust
and University of Birmingham

Sarah R. Beck and Ian J. Mitchell
University of Birmingham

Peter Praamstra
Radboud University

Hardev S. Pall
University of Birmingham and University Hospital Birmingham
NHS Foundation Trust

Objective: Patients with Parkinson's disease (PD) can perform poorly on tasks involving theory of mind (ToM): the ability to reason about mental states. We investigated whether patients' ToM deficits were independent of executive dysfunction. *Method:* Experiment 1 aimed to establish that ToM deficits were present, and 2 following experiments manipulated the working memory (WM) demands of the ToM task. *Results:* In Experiment 1, 15 patients with PD performed significantly more poorly than controls on a false belief vignette task but not on a faux pas task. Errors were related to poor verbal fluency. In Experiment 2, 24 patients with PD made fewer errors on shorter false belief vignettes than the original FBT, and errors on the latter were related to WM impairment. In Experiment 3, the FBT was presented as a comic strip visible throughout questioning, reducing WM demands. Patients ($n = 24$) made memory errors but no false belief errors on the comic strip. They exhibited no verbal fluency or WM impairments, but did exhibit deficits on a black-and-white Stroop task. False belief errors were not correlated with executive performance. *Conclusions:* PD patients made very few ToM errors that were independent of errors on memory questions, so in this sample, ToM deficits per se appear unlikely. However, patients still made errors on ToM tasks when associated incidental WM demands were considerably reduced, highlighting the need for future investigations of ToM in PD to account for the role of more general cognitive restrictions exhibited by even some medicated, early stage patients.

Keywords: Parkinson's disease, executive function, theory of mind, working memory, basal ganglia

In Parkinson's disease (PD), degeneration of dopaminergic neurons in the substantia nigra leads to dysfunction within frontostriatal pathways involved in cognitive functioning. Consequently, patients with PD often exhibit deficits in executive function, performing poorly on tasks requiring planning (e.g., Weintraub et al., 2005) or set shifting (e.g., Lees & Smith, 1983). Working memory (WM) may be particularly affected in PD (e.g., Gabrieli, Singh, Stebbins, & Goetz, 1996) and is needed to remember and manipulate information during a task. Zgaljardic et al. (2006) showed that patients' greatest deficits were apparent on tasks linked to activation of dorsolateral prefrontal cortex (DLPFC), a neural region particularly important for WM (Aleman & van't Wout, 2008).

Patients with PD can also exhibit difficulty with cognitive tasks involving theory of mind (ToM): the ability to reason about mental states (e.g., beliefs, emotions). Studies have indicated that patients with PD are impaired in judging mental states from facial expressions (Suzuki, Hoshino, Shigemasu, & Kawamura, 2006; Yoshimura, Kawamura, Masaoka, & Homma, 2005) or inferring a story character's false belief (Mengelberg & Siegert, 2003). For example, deficits have been reported on the Reading the Mind in the Eyes Test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), which requires participants to judge peoples' mental states from pictures of their eyes (Bodden et al., 2010; Mimura, Oeda, & Kawamura, 2006; Tsuruya, Kobayakawa, & Kawamura, 2011; although another study reported no deficit on this task: Perón et al., 2009). Patients' difficulties can encompass problems with cognitive ToM (e.g., reasoning about beliefs) and affective ToM (e.g., reasoning about emotions), as shown by Bodden et al. (2010). Changes to ToM in PD are important as they can be related to patients' quality of life (Bodden et al., 2010).

Clare M. Eddy, Birmingham and Solihull Mental Health NHS Foundation Trust, Department of Neuropsychiatry, The Barberry, National Centre for Mental Health, Birmingham, United Kingdom, and School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom; Sarah R. Beck and Ian J. Mitchell, School of Psychology, College of Life and Environmental Sciences, University of Birmingham; Peter Praamstra, Department of Neurology, Radboud University, Nijmegen Medical Centre, Nijmegen, The Netherlands; Hardev S. Pall, School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham, and Department of Neurology, Queen Elizabeth Medical Centre, University Hospital Birmingham NHS Foundation Trust, Birmingham.

Correspondence concerning this article should be addressed to Clare M. Eddy, Department of Neuropsychiatry, The Barberry, National Centre for Mental Health, 25 Vincent Drive, Edgbaston, Birmingham B15 2FG, UK. E-mail: clare.eddy@bsmhft.nhs.uk

The neural alterations responsible for ToM deficits in PD are currently poorly understood. Dysfunction of the prefrontal cortex can lead to deficits in ToM (e.g., Channon, Pellijeff, & Rule, 2005; Henry, Phillips, Crawford, Ietswaart, & Summers, 2006; Shamay-Tsoory, Aharon-Peretz, & Levkovitz, 2007; Stone, Baron-Cohen, & Knight, 1998). Specifically, it has been shown that medial orbitofrontal cortex (OFC) is important for ToM (e.g., Fletcher et al., 1995; Gallagher et al., 2000; Stone et al., 1998). One explanation for PD patients' difficulties on ToM tasks is therefore dysfunction in frontostriatal pathways involving medial OFC. However, medial OFC dysfunction may not be necessary for patients to make deficits on ToM tasks. An alternative explanation for deficits in ToM in PD lies in DLPFC dysfunction, as executive deficits resulting from altered functioning of this region can lead to difficulties on ToM tasks. For example, studies of patients with brain injury have shown that impairments in ToM can be linked to deficits in WM (Bibby & McDonald, 2005; Stone et al., 1998) or inhibition (Samson, Apperly, Kathirgamanathan, & Humphreys, 2005), and developmental research has highlighted similar associations (Flynn, 2007; Gordon & Olson, 1998). As difficulties with WM as a result of DLPFC dysfunction are common in PD (e.g., Gilbert, Belleville, Bherer, & Chouinard, 2005), this could help explain patients' deficits on reasoning tasks.

A number of studies provide support the possibility that ToM deficits in PD are a result of executive dysfunction. In addition to showing that WM deficits can lead to poor inferential reasoning in PD (Monetta, Grindrod, & Pell, 2008), Monetta, Grindrod, and Pell (2009) found that patients with PD performed poorly on first- and second-order belief questions during a pragmatic comprehension task, and these difficulties were related to poor performance on measures of WM and verbal fluency. In another study (Saltzman, Strauss, Hunter, & Archibald, 2000), ToM difficulties on a false belief task were related to executive deficits. However, other studies have reported no such link. These include a study conducted by Bodden et al. (2010), which that found patients with PD exhibited most difficulties with second-order ToM. As second-order ToM involves inferring one character's belief about another character's mental state, it is likely to make greater memory and executive demands than first-order ToM, which involves reasoning about one character's mental state. However, Bodden et al. found no evidence that these difficulties with second-order ToM were linked to executive dysfunction.

When executive dysfunction leads to ToM deficits in PD, difficulties with cognitive ToM could be more likely. Perón et al. (2009) reported that advanced but nondemented medicated PD patients (but not early stage PD patients) performed poorly on cognitive ToM questions on a faux pas task, which described one story character making a socially inappropriate remark that could upset or offend another character. These authors suggest that the understanding of cognitive mental states (cognitive ToM) such as intentions or beliefs could be more vulnerable in PD than the understanding of emotions (affective ToM). One study of patients with frontal lesions (Shamay-Tsoory & Aharon Peretz, 2007) showed that impairments in cognitive ToM were more often associated with lesions to DLPFC, which is crucial for many executive tasks. Specific difficulties with cognitive ToM could therefore occur in PD because dysfunction occurs within frontostriatal pathways involving DLPFC leading to executive dysfunction and

cognitive ToM relies more on executive functioning than does affective ToM.

In this study, we conducted three experiments to investigate whether difficulties on verbal ToM tasks exhibited by individuals with PD could be explained by deficits in executive dysfunction. We used both a standard false belief task and a faux pas task to assess both cognitive and affective ToM. Experiments 2 and 3 involved ToM tasks designed to make fewer cognitive demands. These tasks included very short, noninferential false belief tasks and false belief vignettes presented in a comic strip format. We also used executive tasks that included measures of verbal fluency, inhibition, and a WM manipulation span task. A verbal fluency measure was included because although Bodden et al. (2010) reported no differences between patients with PD and controls on a WM span task and the Trail Making Test, these patients did perform more poorly on a verbal fluency measure, suggesting that this task could be sensitive to PD. The inhibitory and WM tasks were selected because studies have highlighted links between ToM performance and inhibition and WM (Carlson, Moses, & Breton, 2002).

Experiment 1

In Experiment 1, we tested ToM in patients with PD using two ToM tasks. We included an executive measure that may be sensitive to PD to determine whether ToM deficits were correlated with executive dysfunction. This experiment provided a basis for later experiments using modified ToM tasks, which were intended to further characterize patients' ToM deficits and their relationship with executive functioning.

The first ToM task was a vignette test devised by Apperly, Samson, Chiavarino, and Humphreys (2004). Each vignette features the unexpected transfer of a person or object while one character is absent, resulting in this character holding a false belief relating to the current location of the object. Participants were required to infer this false belief (cognitive ToM). There were two questions to assess memory, and another question assessed counterfactual thinking, which occurs after an event when a person imagines how things could have turned out differently. Riggs, Peterson, Robinson, and Mitchell (1998) found evidence of a relationship between children's performance on ToM and counterfactual reasoning tasks. This may be because both counterfactual thinking and ToM involve looking at the world from different perspectives. McNamara, Durso, Brown, and Lynch (2003) found evidence for a possible counterfactual deficit in PD.

The other ToM task was a faux pas test (Gregory et al., 2002) that required the identification and understanding of a story character's actions that could have upset another character. The recognition that a faux pas has occurred relies on the attribution of a false belief to the protagonist: that what they say or do will not offend (cognitive ToM), as well as the realization of the victim's feelings of hurt or offense (affective ToM). Kawamura and Koyama (2007) reported unpublished research (Oeda, 2003) that found that although PD patients performed similarly to controls on questions assessing faux pas recognition, they performed significantly more poorly on related false belief questions, despite intact recall of factual information.

Method

Participants. Fifteen volunteers with PD from the Neurosciences Outpatients Department at the Queen Elizabeth Hospital, Birmingham, United Kingdom, participated. The diagnosis of PD had been confirmed by a consultant neurologist with an interest in movement disorders and all patients fulfilled the U.K. PDS Brain Bank criteria (Parts 1 and 2) for diagnosis of PD. They were invited to participate by their consultant neurologist, who ensured that all patients selected neither showed frank symptoms of dementia nor exhibited motor symptoms rated at greater than Stage 3 on the Hoehn and Yahr (1967) scale (i.e., patients had mild to moderate disease with impaired balance, but were physically independent). All patients were taking standard prescribed dopaminergic medications. There were 10 men and five women; mean age was 65.6 years ($SD = 9.60$) and mean education was 14.4 years ($SD = 2.97$). Mean time since disease onset was mean 9.53 years ($SD = 5.85$, range $= 2–19$ years). Ten controls (six men and four women) with a mean age of 60.7 years ($SD = 9.63$) and mean education of 14.3 years ($SD = 1.49$) also participated.

Procedure. The study was approved by South Birmingham NHS Research Ethics Committee. All participants read information sheets about the study and gave written informed consent. The order of administration of the tasks was counterbalanced across participants, although the vignettes within the faux pas and ToM vignette tasks remained in a fixed order. Each vignette was read to the participant twice before they were asked the corresponding questions.

Tasks.

False belief vignette test. Participants were read four unexpected transfer style vignettes. Example: "Andrew is in bed. Susie goes to the shops and while she's away, Andrew goes to the school." Each vignette was read twice and followed by four questions: Counterfactual, "Where would Andrew be if he hadn't gone to the school?" False belief, "Where does Susie think Andrew is?" Memory, "Where was Andrew at the start?" Reality, "Where is Andrew now?" There were two forced-choice responses, the target's original and current locations. Vignettes were presented in a set order, but questions were counterbalanced in pairs (first and second false belief or counterfactual; third and fourth memory or reality).

Faux pas test. Participants were read eight vignettes. Four test vignettes described a potentially offensive faux pas. For example, Jill has moved house and bought new curtains, and Lisa says the curtains are horrible. Four control stories involved no faux pas. Test and control vignettes were presented in a fixed pseudorandom order. The first two questions after a vignette assessed recognition of faux pas, "Did someone say something they shouldn't have?" If yes, "Who was it and what did they say?" If faux pas was identified, two further probes asked, "Why shouldn't they have said that?" and "Why do you think they did say it?" Another question assessed fact recall: "What had Jill just bought?" Finally, desire and belief questions were asked about all test vignettes: "Did Lisa like the curtains?" and "Did Lisa know that Jill had chosen the curtains?"

FAS verbal fluency test. Participants were asked to say as many words as they could think of beginning with a given letter of the alphabet until being asked to stop, apart from proper names.

One minute was given for the letters *F*, *A*, and then *S*. Total scores were calculated.

Results

Nonparametric Mann–Whitney *U* tests and Spearman's rho correlation coefficients were calculated as data were not normally distributed.

Patients and controls did not differ significantly for age ($U = 50.5$, $p = .177$) or education ($U = 56.0$, $p = .311$).

Group comparisons.

FAS verbal fluency test. Patients generated fewer words ($M = 40.93$, $SD = 18.29$) than controls ($M = 55.08$, $SD = 15.65$) over the FAS test, a difference that was significant ($U = 37.0$, $p = .035$).

Faux pas test. Patients did not perform significantly differently from controls when asked to recognize faux pas ($U = 59.0$, $p = .313$). Patients ($n = 15$) made 14 errors in total ($M = 0.93$, $SD = 1.28$, $Mdn = 0.0$, range $= 0–4$), and controls ($n = 10$) made four errors ($M = 0.4$, $SD = 0.70$, $Mdn = 0.0$, range $= 0–2$). Errors were made on both test and control vignettes. However, patients did exhibit poorer performance on fact check questions ($U = 37.5$, $p = .027$) and a trend for poorer performance on the desire question ($U = 47.0$, $p = .057$). There was no difference for the belief question ($U = 53.5$, $p = .186$).

False belief vignette test. Patients made an average of 1.46 errors on the false belief vignette test and performed significantly more poorly than controls ($U = 56.0$, $p = .001$), who performed at ceiling. The number of errors made on memory, false belief, and counterfactual questions made by patients with PD did not differ significantly, $\chi^2(3) = 3.00$, $p = .392$. Of the seven false belief errors made by patients (see Table 1), only three were not accompanied by errors on memory check questions.

Correlations. Correlations were conducted to examine relationships between performance on executive and ToM tasks and time since disease onset and task performance. A significant correlation was found between patient performance on the FAS test and false belief vignette test ($r_s = -.529$, $p = .043$). Better executive performance as measured by high scores on the FAS test was associated with fewer errors on the ToM test. This correlation was not significant when using Bonferroni correction for multiple comparisons. No other significant correlations were found.

Discussion

Patients with PD performed significantly more poorly than controls on the first-order false belief vignette test, which could indicate a ToM deficit. False belief errors found in the present study are in accordance with other reports of false belief deficits in PD (e.g., Mengelberg & Siegert, 2003). However, false belief errors were usually accompanied by errors on memory and reality questions, so the most likely explanation is that ToM performance was affected by memory impairment. PD was also associated with poorer performance on the verbal fluency task, and the correlation between performance on this measure and the false belief vignette test could suggest that executive dysfunction contributed to errors on the vignette test.

Patients made very few errors on counterfactual questions, providing no evidence for a counterfactual deficit, in contrast to the results of McNamara et al. (2003).

Table 1
Errors Made by Patients With Parkinson's Disease on the False Belief Vignette Test in Experiments 1 and 2

Experiment	False belief	Counterfactual	Memory	Reality	Total
1 (*n* = 15)					
Errors/total possible	7/60	3/60	6/60	6/60	22/240
Percentage incorrect	11.67	5	10	10	9.16
SD	0.64	0.41	0.63	0.83	
Number of patients making errors	6	3	5	4	8
2 (*n* = 24)					
Errors/total possible	9/96	14/96	11/96	13/96	47/384
Percentage incorrect	9.38	14.58	11.46	13.54	12.24
SD	0.58	1.14	0.51	0.59	
Number of patients making errors	8	7	11	12	15

The possibility that patients' errors on the false belief vignette test reflected a specific deficit in ToM is further weakened by their performance on the faux pas test. It is interesting that patients performed well on the faux pas task despite false belief errors on the vignette test. One explanation for this difference is that semantic social script knowledge about appropriate behavior could be used to answer this question correctly without a need to rely on ToM. Thus, the task required less online reasoning than the false belief task, which contained more arbitrary information about social interaction. Patients could therefore answer the faux pas recognition question well in spite of a ToM deficit, or indeed executive difficulties, if relevant semantic knowledge about social norms was intact.

An alternative explanation for better performance on the faux pas task than the vignette task is linked to the involvement of affective ToM in the faux pas task. The discrepancy in performance across these tasks could suggest that patients were more readily able to reason about the emotional response of the victim in the faux pas scenarios than cognitive ToM involving false belief. It could be that Perón et al. (2009) found a deficit on the faux pas task because the later stage patients tested in their study did have a deficit in affective ToM, perhaps in association with greater medial prefrontal dysfunction. Whatever the case, neither of these possibilities can easily explain why even despite making significantly more errors on fact questions during the faux pas task, patients with PD in the current study did not exhibit difficulties with faux pas false belief questions, which would be expected if cognitive ToM difficulties were present in the patient group.

Another difference between performance on the false belief vignette test and faux pas tasks was that errors on the vignette test were linked to poor executive performance, but faux pas performance was not. One explanation for this is the use of stored knowledge rather than online reasoning in the faux pas task, which would mean fewer executive demands. However, it can be speculated that the lack of affective ToM in the false belief vignette test that involved purely cognitive mental states could help explain this differentiation. Perhaps reasoning about more abstract mental states or cognitive ToM is more closely tied to executive function. Studies have shown that tasks involving cognitive ToM appear to more often recruit DLPFC regions important for executive abilities (Shamay-Tsoory & Aharon-Peretz, 2007).

Experiment 1 showed that patients with PD can exhibit false belief errors on some tasks and executive difficulties suggestive of prefrontal dysfunction. Further work is needed to determine whether these false belief errors are purely a result of executive dysfunction. The number of memory errors made by patients on the ToM vignettes was surprising, as the vignettes were fairly short and control performance was at ceiling. Experiment 2 investigated the relationship between WM and false belief errors.

Experiment 2

The pattern of errors seen in Experiment 1 could reflect WM deficits, which are commonly reported in PD (e.g., Gabrieli et al., 1996) and are likely to be linked to DLPFC dysfunction (Aleman & van't Wout, 2008). Indeed, the finding of a deficit in verbal fluency could be suggestive of DLPFC dysfunction, as fluency tasks also activate DLPFC (Gaillard et al., 2000). Experiment 2 was designed to investigate whether the errors made by patients could indeed be attributed to deficits in WM.

One way in which WM demands were reduced in Experiment 2 was by decreasing the amount of information in the vignettes. The shortest new vignettes were two lines long and were of the "unexpected transfer" style used in the false belief vignette test, which were six to seven lines long. Another way that the potential impact of WM impairment was reduced was through the inclusion of noninferential "deceptive box" style ToM vignettes. Inferential reasoning involves using available information to deduce a likelihood that is not explicitly stated. The deceptive box vignettes described one character being deceived by another about the true identity of an object. The fact that a mistaken belief was held by the deceived character was explicitly stated: Participants were told that the character in question believed the lie the other character told them. This task therefore involved thinking about the characters' mental states but inference beyond the information presented was not necessary. Inferential reasoning deficits have been reported in PD. For example, Berg, Bjornram, Hartelius, Laakso, and Johnels (2003) found that patients with PD exhibited poor inferential reasoning about story characters during linguistic tasks, and Monetta et al. (2008) showed that patients with PD can exhibit deficits in inference generation as a result of WM dysfunction. One specific link between WM impairment in PD and ToM deficits could therefore be through impaired inferential reasoning.

The patients tested in this experiment completed the FAS test and a WM test, the Adaptive Digit Ordering Test (DOT-A: Werheid et al., 2002; adapted from Cooper, Sagar, Jordan, Harvey, & Sullivan, 1991). This WM test assesses manipulation span, which may be particularly vulnerable to PD (Werheid et al., 2002).

Method

Participants. Twenty-four new volunteers with PD (18 men and six women) participated in this experiment. They were selected based on the criteria described in Experiment 1 and all were taking standard dopaminergic treatments. Their mean age was 66 years ($SD = 8.99$, range = 49–86 years) and they had a mean of 12.83 years ($SD = 2 20$, range = 11–19 years) of education. Mean time since disease onset was 8.02 years ($SD = 4.94$, range = 2–20 years). The main interest of Experiments 2 and 3 were within-group analyses, so controls were used only for additional executive measures. Twenty-eight new controls (14 men and 14 women, mean age = 60.38 years, $SD = 10.33$, range = 41–78 years; mean education = 12.75 years, $SD = 2.03$, range = 11–17 years) were recruited to provide comparative data for the DOT-A.

Procedure. The procedure for this experiment was the same as that for Experiment 1.

Tasks.

ToM vignettes. There were three types of ToM vignettes, presented as a block in a fixed order. The order of presentation was counterbalanced by type, so that vignettes of the three types were distributed as evenly as possible throughout the fixed order.

Short, noninferential "deceptive box" style vignettes. Each deceptive box vignette described one character telling another character a lie to deceive them about the identity of an item. Example: "It is Arthur and Beryl's anniversary. Beryl tells Arthur that she has bought him a CD. In fact, she has bought him tickets to a concert. Arthur doesn't know that Beryl is going to surprise him, so he believes Beryl." These vignettes were followed by false belief (e.g., "What does Arthur think Beryl has bought him for his birthday?") and reality ("What has Beryl really bought for Arthur?") questions, with two forced-choice options.

Short, "unexpected transfer" vignettes. The unexpected transfer vignettes had a similar structure to those in the false belief vignette test but were shorter. Example: "Judy leaves some cigarettes on the coffee table. While she is in bed asleep, her son takes the cigarettes and puts them in his schoolbag." Participants were asked false belief (e.g., "Where does Judy think the cigarettes are?") and reality (e.g., "Where are the cigarettes really?") questions, and two forced-choice options were given.

False belief vignette task. The false belief vignette task was the same as that in Experiment 1.

FAS test. The FAS test was the same as that in Experiment 1.

DOT-A. This task assesses verbal WM manipulation span. Participants were asked to listen to individual streams of digits, with the length of the stream increasing from three to eight digits as testing progressed. Participants were asked to remember the digits and say them back in ascending order immediately after presentation. A participant's maximum span was the longest stream responded to correctly. Half a point was deducted from the maximum span of a participant if only one stream of that length was responded to correctly.

Results

Patients' education did not differ from that of controls tested on the DOT-A in this experiment, $t(50) = 0.142$, $p = .888$, but age was slightly higher than controls, $t(50) = 2.043$, $p = .042$. Age and education for these patients were not significantly different from

that of controls tested on the FAS test in Experiment 1: age, $t(37) = 1.35$, $p = .186$; education, $t(37) = -1.40$, $p = .171$.

ToM tasks.

"Deceptive box" style vignettes. A total of 18 errors were made on the deceptive box vignettes, with a mean of 0.75 per patient ($SD = 1.07$, $Mdn = 0.0$, range = 0–4). An equal number of false belief and reality errors were evident, and 10 patients made at least one error on the task.

"Unexpected transfer" vignettes. Patients made fewest errors (12) on unexpected transfer vignettes, leading to a mean of 0.5 errors per patient ($SD = 1.25$, $Mdn = 0.0$, range = 0–5). Patients made an equal number of false belief and reality errors on these vignettes. Only four patients made errors on this task.

False belief vignette task. Overall, patients made 47 errors on this test (see Table 1), and each patient made an average of 1.96 errors ($SD = 2.07$, $Mdn = 1.5$, range = 0–7), of a possible 16 (Experiment 1 controls exhibited scores at ceiling). A Friedman test indicated no significant difference in performance on false belief, counterfactual, memory, and reality questions, $\chi^2(3) = 1.41$, $p = .703$. Of the nine false belief errors made by patients (see Table 1), only four were not accompanied by errors on memory check questions.

Comparison between vignette types. As the false belief vignette task had more questions than the deceptive box and unexpected transfer vignettes, error counts were converted into percentages. Wilcoxon signed-ranks test indicated that significantly more errors were made on the false belief vignette test than on the short unexpected transfer vignettes, $z = -3.42$, $p = .001$, and deceptive box vignettes, $z = -2.16$, $p = .031$. The number of errors made on deceptive box and unexpected transfer vignettes did not differ significantly, $z = -0.998$, $p = .318$.

Executive tasks.

FAS test. The mean number of words generated by patients on the FAS test was 38.83 ($SD = 17.28$, $Mdn = 40.5$, range = 14–95), which was significantly lower than the controls tested in Experiment 1, who achieved a mean score of 54.40 words ($SD = 16.0$, $Mdn = 47.0$, range = 30–95), $U = 102.5$, $p = .025$.

DOT-A. Patients achieved a mean WM manipulation span of 4.58 digits ($SD = 0.99$, $Mdn = 4.5$, range = 2.5–6.5), which was significantly lower than controls' mean span of 5.91 digits ($SD = 0.65$, $Mdn = 5.75$, range = 4.5–7.0), $U = 585$, $p < .001$.

Correlations. Correlations were conducted to examine relationships between patient performance on executive tasks and ToM vignette performance and time since disease onset and task performance (time since disease onset could be related to task performance as PD involves progressive neurodegeneration). DOT-A WM scores correlated significantly with performance on the longer ToM vignettes ($r_s = -.413$, $p = .045$), indicating that more errors were made by patients with poorer WM. No other significant correlations were evident. The aforementioned correlation was not significant when using Bonferroni correction for multiple comparisons.

Discussion

Patients exhibited frontal executive dysfunction as evidenced through poorer performance on both the FAS test and DOT-A. Nevertheless, performance on the false belief vignette test was related to WM performance, but not to performance on the FAS

test, in contrast to Experiment 1. This inconsistency in findings probably indicates that the relationship between ToM and verbal fluency is generally weak, and other measures of executive function may be more closely related to performance. Bodden et al. (2010) reported deficits on ToM tasks involving second-order ToM that did not appear to be linked to WM deficits in early stage patients. However, Bodden et al. used a measure of simple span rather than manipulation span, and the latter may be more sensitive to WM impairment in PD (Werheid et al., 2002). One limitation of Experiment 2 is that the PD patients tested were slightly older than controls who completed the DOT-A.

As WM scores correlated with performance only on longer vignettes (the false belief vignette test), it is likely that it was the amount of information contained in the task that was closely related to WM difficulties and patients' errors. Patients made fewest errors on the unexpected transfer vignettes, the shortest vignettes, even though they contained the same type of false belief reasoning as the false belief vignette test. Reducing some information processing demands can therefore improve patients' performance on ToM tasks. However, patients also made some errors on deceptive box vignettes, which involved thinking about mental states but did not require inferential reasoning. Although these errors could be more likely to reflect specific difficulties with reasoning about mental states, the presence of memory errors on this task means that an executive basis for patients' deficits cannot be ruled out.

Experiment 2 demonstrated that reducing the amount of information contained in ToM vignettes (and so probably WM demands) can improve PD patients' performance. However, false belief and memory errors were not completely eliminated for very short and noninferential ToM vignettes. Experiment 3 investigated whether reducing memory demands even further could reveal intact ToM performance, which would suggest that patients' difficulties on these tasks are not a result of ToM difficulties per se.

Experiment 3

Experiment 3 involved the investigation of patients' performance on the false belief vignette test when WM demands incidental to the task were reduced. The false belief vignette test was presented to patients in three different formats with varying WM demands. For Condition A, false belief vignettes were simply read to participants, as in previous studies. For Condition B, vignettes were also presented with transient additional visual information, in the form of a flipbook of cartoons, with accompanying story text. For Condition C, cartoons and story text were used again, but were presented in the form of a comic strip, which was present throughout questioning. As well as reducing WM load, presenting information in written as well as verbal form increases participants' attention (Bibby & McDonald, 2005). It was hypothesized that patients' performance would be best in Condition C. Ceiling performance on false belief questions in this condition could indicate intact ToM.

Three executive tasks were used to investigate relationships between ToM and executive performance, the FAS test, the DOT-A, and a black-and-white Stroop, to assess inhibition. Like WM, inhibition has been linked to ToM performance (e.g., Carlson & Moses, 2001; Samson et al., 2005). Developmental research has shown that the black-and-white Stroop is a useful measure of inhibition (e.g., Beck, Riggs, & Gorniak, 2009), and a previous study revealed a deficit on this task in individuals with another movement disorder (Eddy, Mitchell, Beck, Cavanna, & Rickards, 2010). The black-and-white Stroop may be preferable to a traditional Stroop test when testing patients with movement disorders, as motor symptoms may be more likely to interfere with performance on a traditional Stroop because of the size and arrangement of stimuli.

Method

Participants. Twenty-four outpatients with PD from the Neurosciences Department at the Queen Elizabeth Hospital, Birmingham, took part in this experiment. They were selected on the basis of the criteria described in the procedure of Experiment 1 and were all taking prescribed dopaminergic medications. There were 15 men and nine women, of mean age 62.45 years ($SD = 9.81$, range $= 41–75$ years), with a mean of 13.25 years ($SD = 2.23$, range $= 11–17$ years) of education. Mean time since disease onset was 6.9 years ($SD = 4.80$, range $= 1–17$ years). Sixteen controls (eight women) completed the Stroop test. Their mean age was 56.63 years ($SD = 6.92$, range $= 41–71$ years), and their mean was 12.4 years ($SD = 1.71$, range $= 11–16$ years).

Procedure. The procedure was the same as that in Experiment 1.

Tasks.

False belief vignette test with three conditions. A longer form of the false belief vignette test was used, consisting of 12 vignettes taken from Apperly et al. (2004; see Experiment 1). Each participant received four of the 12 vignettes in Conditions A, B, and C. The order of conditions was counterbalanced across participants, and each vignette appeared equally often in each condition overall.

In Condition A, the vignettes were read out loud twice to each participant, with no accompanying visual representation.

In Condition B, the vignettes were read twice with an accompanying flipbook. Each flipbook had four pages and featured a cartoon with corresponding story text underneath it. The first cartoon illustrated the object's starting position. The second illustrated the absence of the character holding the false belief. The third picture showed an event that led to the change in location of the object, and the final picture showed the object's new location. The appropriate flipbook was read twice to the participant, and was removed before questioning.

For Condition C, the same pictures and text used for flip books were presented in the form of a comic strip, with four pictures in a single row (see Figure 1). This was displayed as the vignette was read twice to the participant and was visible throughout questioning.

FAS test. The FAS test was the same as that in Experiment 1.

DOT-A. The DOT-A was the same as that in Experiment 2.

Black-and-white Stroop test. The black-and-white Stroop test assesses inhibitory skills in a very similar way as a traditional Stroop task by requiring the participant to suppress a prepotent response to say the name of one color and instead say the name of a different color. Stimuli consisted of a page of 40 equally sized squares colored black or white and arranged in a pseudorandom order. For the baseline condition, participants were asked to say the color of each square, going across each row from left to right, and not to correct any incorrect answers. For the test condition, participants were told to say *black* if they saw a white square and
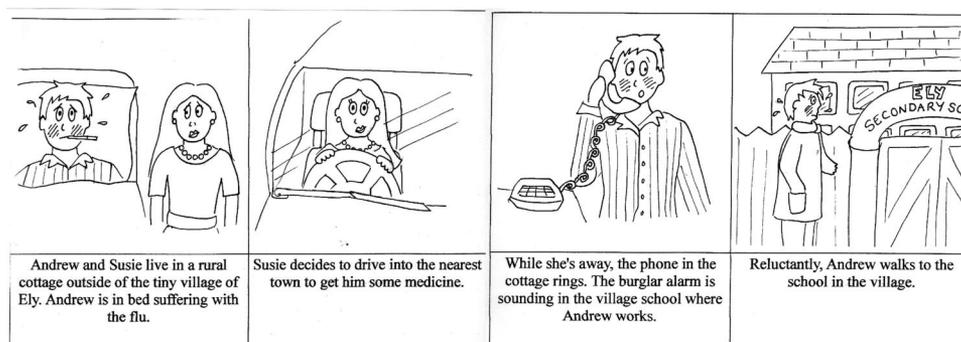
Figure 1. Example of a false belief vignette comic strip used in Experiment 3 (vignette text from Apperly et al., 2004).

say *white* for a black square. Time taken and number of errors were recorded. This type of Stroop may be preferable to a traditional Stroop when working with patients with movement disorders as it contains larger and fewer stimuli, making performance less susceptible to problems such as medication-induced dyskinesia and eye-movement difficulties in scanning the array.

Results

Patients did not significantly differ in age or years of education compared with controls tested in Experiment 1: age, $t(37) = 0.070$, $p = .944$; education, $t(37) = -0.790$, $p = .435$, or Experiment 2: age, $t(50) = 0.748$ $p = .458$; education, $t(50) = 1.166$, $p = .249$, but age was marginally significantly different for Experiment 3: age, $t(38) = 2.058$, $p = .046$; education, $t(38) = -1.520$, $p = .137$.

ToM false belief vignettes. Ten of the 24 patients made at least one error over the whole vignette test (see Table 2). Seven made errors on the verbal condition, five made errors during the flipbook condition, and four made errors when responding to comic strip questions. A mean of 0.38 errors was made by each patient during the flipbook ($SD = 0.9$, $Mdn = 0.0$, range = 0–3) and comic strip ($SD = 0.9$, $Mdn = 0.0$, range = 0–3) conditions. Slightly more errors ($M = 0.54$, $SD = 1.1$, $Mdn = 0.0$, range = 0–4) were made in response to verbal-only presentation. However,

a Friedman test, $\chi^2(2) = 1.27$, $p = .529$, indicated no significant difference in the number of errors made by patients across presentation conditions.

There was little difference in the number of false belief, counterfactual, memory, and reality errors made within each condition. Only four false belief errors were made by PD patients during this experiment. Three of these were not accompanied by errors on memory and reality questions. However, no false belief errors were made in the comic strip condition. When the total number of false belief, counterfactual, memory, and reality errors were compared after being collapsed across conditions, a Friedman test failed to reveal any significant differences, $\chi^2(3) = 5.59$, $p = .134$.

Executive tasks. Patient performance on the executive tasks was compared with the performance of controls tested in Experiments 1, 2, and 3.

FAS test. Patients generated a mean of 46.29 words ($SD = 15.0$, $Mdn = 49.0$, range = 20–78) on this task, which did not differ significantly from the mean 54.4 words ($SD = 16.0$, $Mdn = 47.0$, range = 30–95) produced by controls in Experiment 1, $U = 146.0$, $p = .326$.

DOT-A. The mean WM manipulation span for the current patient group was 5.60 digits ($SD = 1.2$, $Mdn = 5.5$, range = 4–8). This did not differ significantly from the control mean of

Table 2
Errors Made by Patients With Parkinson's Disease (n = 24) on the False Belief Vignette Test in Experiment 3

Condition	False belief	Counterfactual	Memory	Reality	Total
Verbal					
Errors/total possible	3/96	2/96	3/96	5/96	13/384
Percentage incorrect	3.13	2.08	3.13	5.21	3.39
SD	0.40	0.30	0.45	0.40	
Number of patients making errors	2	2	2	5	7
Flipbook					
Errors/total possible	1/96	1/96	3/96	4/96	9/384
Percentage incorrect	1.04	1.04	3.13	4.17	2.34
SD	0.20	0.20	0.40	0.50	
Number of patients making errors	1	1	2	3	5
Comic strip					
Errors/total possible	0/96	4/96	2/96	3/96	9/384
Percentage incorrect	0	4.17	2.08	3.13	2.34
SD	0	0.60	0.30	0.40	
Number of patients making errors	0	2	2	2	4

5.91 digits ($SD = 0.65$, $Mdn = 5.75$, range $= 4.5–7.0$) reported in Experiment 2, $U = 396.0$, $p = .258$.

As WM span ranged from four to eight digits across the patient group, the group was divided according to scores on this measure. Patients were divided into two groups on the basis of comparison of each participant's span with the group median of 5.5. Two patients had scores equal to this median value, so these patients were excluded. The WM+ group (scores above 5.5) and WM– group (scores below 5.5) each contained 11 patients.

Patients in the WM– group exhibited lower scores on the FAS test compared with the WM+ group, $U = 28$, $p = .033$. Patients in the WM+ group exhibited a lower mean number of errors on the black-and-white Stroop, $U = 24.0$, $p = .013$. However, there was no significant difference for Stroop times, $U = 45.0$, $p = .309$.

The WM– group contained seven patients who made a total of 19 errors on the false belief vignette test (verbal: 11 errors; flip book: five errors; comic strip: three errors). The WM+ group contained three patients who made a total of 12 errors on the false belief vignette test (verbal: two errors; flip book: four errors; comic strip: six errors). The overall number of errors made by the two groups did not differ significantly, $U = 41.5$, $p = .171$. The number of errors made across conditions did not differ significantly for the WM+, $\chi^2(2) = 1.4$, $p = .497$, or WM–, $\chi^2(2) = 4.52$, $p = .104$, group.

Black-and-white Stroop test. Patients with PD made significantly more errors on the black-and-white Stroop compared with controls, $U = 109.5$, $p = .013$. Patients ($n = 24$) made a total of 44 errors ($M = 1.83$ per patient, $SD = 2.35$, $Mdn = 1.0$, range $= 0–8$) and the control group ($n = 16$) made just eight errors in total ($M = 0.5$ per control, $SD = 1.26$, $Mdn = 0$, range $= 0–5$). However, patients did not show a greater effect of interference in terms of time taken, $U = 140.5$, $p = .155$. Patients' mean time difference of 10.74 s ($SD = 8.37$, $Mdn = 9.00$, range $= 0.86–36.29$ s) was not significantly greater than controls' mean time difference of 8.20 s ($SD = 7.41$, $Mdn = 7.14$, range $= 0.41–33.37$ s).

Correlations. Analyses were conducted to identify possible relationships between patient performance on executive tasks and ToM vignette performance and time since disease onset and task performance. Significant relationships were found between FAS score and DOT-A score ($r_s = .598$, $p = .002$), and between the number of errors made on the Stroop test condition and DOT-A score ($r_s = -.531$, $p = .008$). No other significant correlations were found.

Discussion

Overall, patients with PD seemed to make relatively fewer errors on the false belief vignette test than in Experiments 1 and 2. It is possible that the new presentation conditions contributed to this finding. However, the task could also have elicited fewer errors because patients in Experiment 3 exhibited no significant deficits in verbal fluency or WM, and their WM scores appeared higher than the PD patients tested in Experiment 2. In PD, if executive deficits are likely to contribute to errors on the false belief vignette test, this could explain why patients in Experiment 3 made fewer errors than patients tested in earlier experiments. One finding that could weaken the above proposal is that in Experiment 3, performance on the ToM vignette task was not related to executive performance on tasks assessing WM, fluency, or inhibition.

The finding that there was no relationship between ToM vignette errors and WM could have arisen because of the format of presentation of the ToM vignette task used in this experiment. The task as a whole should have made fewer demands on WM resources. An alternative explanation is that not all of the errors made by patients were due to deficits in WM. This proposal is supported by the observation that errors were made on ToM vignettes in Experiment 3 despite the finding that the patient group did not exhibit significantly lower WM scores than controls. Moreover, no significant differences in performance on the ToM vignette test were seen when patients were divided into two groups containing patients with high or low WM scores.

In summary, this group of patients with fewer executive deficits made fewer errors on the false belief vignette test overall and very few false belief errors. No false belief errors were made in the condition designed to most reduce WM demands (comic strip), but errors were not correlated with WM dysfunction.

General Discussion

We investigated whether deficits in ToM in PD made on standard verbal false belief tasks could be independent of executive dysfunction. The errors made on the false belief vignette test in Experiments 1 to 3 could reflect a deficit in ToM per se, the impact of WM limitations on task performance, or less specific cognitive difficulties. The possibility that deficits were due to difficulties in ToM per se is weakened by the fact that errors were not specific to false belief questions and were also apparent on counterfactual and memory check questions.

The likelihood that patients' errors reflected a ToM deficit per se is weakened by the finding that patients did not perform significantly more poorly than controls on the faux pas task. One reason this is surprising is because this task is passed by older children than simpler false belief tasks, and may involve second-order ToM (e.g., Baron-Cohen et al., 1999). Patients demonstrated good understanding of the victim's emotional response to the protagonist's offensive remark (affective ToM) during this task. A deficit on the false belief vignette test but not the faux pas task could imply that patients with PD exhibit specific difficulties in cognitive ToM. However, patients did not perform worse than controls on faux pas false belief questions. This finding does not suggest that patients tested exhibited fundamental deficits in reasoning about cognitive mental states.

The finding that patients with PD performed poorly on the false belief vignette test but not on the faux pas task could be in keeping with the proposal that patients' errors reflected executive impairments, rather than a deficit in ToM per se, if the faux pas task made fewer executive demands. This may perhaps seem unlikely, as faux pas tasks are passed later in development. However, another explanation for patients' better performance on the faux pas task is that they were able to use intact knowledge about social norms to answer faux pas questions correctly. The use of such knowledge could help avoid specific ToM requirements or related executive demands made by the task.

Studies have shown that like many executive functions, cognitive ToM is likely to rely on DLPFC, and damage to medial prefrontal regions may specifically impair affective ToM

(Shamay-Tsoory & Aharon-Peretz, 2007). Bodden et al. (2010) suggested that early substantia nigra degeneration in PD leads to dysfunction within frontostriatal pathways involving DLPFC first, so cognitive ToM is likely to be more vulnerable in early PD than affective ToM. This could help explain why patients had no difficulty in reasoning about the mental state of the offended victim during the faux pas task.

On many occasions, patients' errors on the false belief vignette test are likely to have reflected executive dysfunction. Deficits in WM were linked to errors on the false belief vignette test in Experiment 2. One way in which ToM deficits may result from WM dysfunction is through poor retention of information in the story in order to infer a character's mental state. When memory demands were reduced by presenting the false belief vignette test as a comic strip left visible throughout questioning (Experiment 3), no false belief errors were made. However, it seems unlikely that this kind of WM demand could explain false belief errors made on very short or noninferential ToM vignettes (Experiment 2).

One strength of this study was the use within-task control trials and manipulation of ToM tasks experiments in an attempt to isolate the contribution of executive demands. However, this was only partly achievable. We were able to show that by reducing the incidental executive demands associated with a false belief task, we could reduce the number of errors made by patients. We have therefore provided evidence that executive difficulties can certainly contribute to ToM errors in PD. However, there is a fundamental difficulty in assessing the contribution of executive difficulties to patients' difficulties when one considers the likely integral role of executive abilities in performing ToM reasoning. For example, there may be an integral role for WM in ToM, whereby this executive function provides a basis for the simultaneous consideration of and manipulation of perspectives (e.g., one's own true belief and a story character's false belief).

As WM difficulties were not always associated with poor performance (e.g., Experiment 3), other factors could have contributed to errors on the ToM tasks. It may be that difficulties remembering information to complete the tasks reflected problems with focusing on relevant information in order to encode and store it. Retrieval deficits are perhaps less likely as patients made errors on questions when given two forced-choice options. However, Faglioni, Botti, Scarpa, Ferrari, and Saetti (1997) found that patients with PD exhibited impairments in story recall and a peculiar pattern of forgetting, indicating that retaining processes could be abnormal.

Reducing the amount of task information may have improved patients' performance by reducing the chances of cognitive resource overload. Dysfunction of DLPFC has been associated with poor attention and arousal, and may result in "nonspecific global cognitive impairment" in PD (Berry, Nicolson, Foster, Behrmann, & Sagar, 1999). At times when cognitive resources are insufficient, patients show blanking and mind wandering (Watts, McLeod, & Morris, 1988). Difficulties with attention and concentration could affect patients' performance on both WM and ToM tasks, leading to unspecified errors.

Patients in the current experiment were always questioned immediately after the vignette was read to them. Some studies suggest that whereas patients perform poorly on memory tasks if tested immediately after presentation, performance can be better after a delay (Knoke, Taylor, & Saint-Cyr, 1998; Sagar, Sullivan,

Gabrieli, Corkin, & Growdon, 1988), indicating that patients with PD have most difficulties soon after registration of information. Such findings could be linked to bradyphrenia or slower encoding processes in PD. Future research should investigate whether patients perform better when questioned after a delay allowing more time for memory consolidation.

An alternative explanation for patients' performance on the ToM tasks is that errors reflect attentional set-shifting deficits. Cronin-Golomb, Corkin, and Growdon (1994) reported that despite intact logical reasoning, patients with PD exhibited deficits on deductive reasoning tasks, which correlated with set-shifting deficits on the Wisconsin Card Sorting Test. It may be that processes involved in set-shifting tasks are linked to ToM because of a role for switching between perspectives in order to adopt another person's point of view and so understand their mental state. Further studies of ToM in PD should include set-shifting measures.

There are a number of limitations associated with the current study. One of these is that we did not include nonverbal measures of ToM, which would have provided a useful comparison and could have the potential to be less influence by executive problems. Although we examined relationships between disease progression and task performance, it would have been preferable to use Unified Parkinson's Disease Rating Scales to assess the possible influence of disease severity. Another important limitation is the use of different patients across experiments. The superior executive performance exhibited by patients tested in Experiment 3 compared with Experiments 1 and 2 emphasizes the heterogeneity within this patient population, which makes it difficult to compare the findings across experiments. It may be useful for future studies to compare the ToM performance of patients who have only recently developed the symptoms of PD with patients who have had the condition for a number of years. Finally, a further important point to consider when assessing cognitive performance in PD populations is the potential influence of dopaminergic medications. For example, it has been suggested that elevated dopamine levels could be linked ToM deficits (Abu-Akel, 2003) and such relationships would be pertinent to PD. It is therefore important to consider how much medication-related factors could account for findings. This limitation applies to the current study, as all patients tested were taking these medications.

In conclusion, our findings demonstrate that patients with PD can find it difficult to infer false belief. Indeed, patients with early to moderate PD can be challenged by ToM tasks, especially if the verbal memory demands are high. Perhaps surprisingly, errors can remain when incidental WM demands of verbal ToM tasks are reduced and patients are required to process very little information. We cannot be sure that all of the errors in this study were independent of executive deficits, so we cannot assume they all represent deficits in ToM per se. To provide convincing evidence that patients with PD do exhibit deficits in ToM per se, future investigations will need to take into account the extent of cognitive restrictions exhibited by even some medicated, early stage patients, which will exert a critical impact on ToM performance.

References

Abu-Akel, A. (2003). The neurochemical hypothesis of "theory of mind." *Medical Hypotheses, 60,* 382–386. doi: 10.1016/S0306-9877(02)00406-1

Aleman, A., & van 't Wout, M. (2008). Repetitive transcranial stimulation over the right dorsolateral prefrontal cortex disrupts digit span task performance. *Neuropsychobiology, 57,* 44–48. doi:10.1159/000129666

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: Neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience, 16,* 1773–1784.

Baron-Cohen, S., O'Riodan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with autistic spectrum or high functioning autism. *Journal of Autism and Developmental Disorders, 29,* 407–418. doi:10.1023/A:1023035012436

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test Revised Version: A study with normal adults, adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42,* 241–251. doi:10.1111/1469-7610.00715

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2009). Relating developments in children's counterfactual thinking and executive functions. *Thinking and Reasoning, 15,* 3337–3354.

Berg, E., Bjornram, C., Hartelius, L., Laakso, K., & Johnels, B. (2003). High-level language difficulties in Parkinson's disease. *Clinical Linguistics & Phonetics, 17,* 63–80. doi:10.1080/0269920021000055540

Berry, E. L., Nicolson, R. I., Foster, J. K., Behrmann, M., & Sagar, H. J. (1999). Slowing of reaction time in Parkinson's disease: The involvement of the frontal lobes. *Neuropsychologia, 37,* 787–795. doi:10.1016/S0028-3932(98)00137-7

Bibby, H., & McDonald, S. (2005). Theory of mind after traumatic brain injury. *Neuropsychologia, 43,* 99–114. doi:10.1016/j.neuropsychologia.2004.04.027

Bodden, M. E., Mollenhauer, B., Trenkwalder, C., Cabanel, N., Eggert, K. M., Unger, M. M., . . . Kalbe, E. (2010). Affective and cognitive theory of mind in patients with Parkinson's disease. *Parkinsonism & Related Disorders, 16,* 466–470. doi:10.1016/j.parkreldis.2010.04.014

Carlson, S. M., & Moses, L. J. (2001). Individual difference in inhibitory control and children's theory of mind. *Child Development, 72,* 1032–1053. doi:10.1111/1467-8624.00333

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11,* 73–92. doi:10.1002/icd.298

Channon, S., Pellijeff, A., & Rule, A. (2005). Social cognition after head injury: Sarcasm and theory of mind. *Brain and Language, 93,* 123–134. doi:10.1016/j.bandl.2004.09.002

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

Cronin-Golomb, A., Corkin, S., & Growdon, J. H. (1994). Impaired problem solving in Parkinson's disease: Impact of a set-shifting deficit. *Neuropsychologia, 32,* 579–593. doi:10.1016/0028-3932(94)90146-5

Eddy, C. M., Mitchell, I. J., Beck, S. R., Cavanna, A. E., & Rickards, H. E. (2010). Impaired comprehension of nonliteral language in Tourette syndrome. *Cognitive and Behavioral Neurology, 23,* 178–184.

Faglioni, P., Botti, C., Scarpa, M., Ferrari, V., & Saetti, M. C. (1997). Learning and forgetting processes in Parkinson's disease: A model based approach to disentangling storage, retention and retrieval contributions. *Neuropsychologia, 35,* 767–779. doi:10.1016/S0028-3932(96)00125-X

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition, 57,* 109–128. doi:10.1016/0010-0277(95)00692-R

Flynn, E. (2007). The role of inhibitory control in false belief understanding. *Infant and Child Development, 16,* 53–69. doi:10.1002/icd.500

Gabrieli, J. D. E., Singh, J., Stebbins, G. T., & Goetz, C. G. (1996). Reduced working memory span in Parkinson's disease: Evidence for the role of a frontostriatal system in working and strategic memory. *Neuropsychology, 10,* 322–332. doi:10.1037/0894-4105.10.3.321

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia, 38,* 11–21. doi:10.1016/S0028-3932(99)00053-6

Gaillard, W. D., Hertz-Pannier, L., Mott, S. H., Barnett, A. S., LeBihan, D., & Theodore, W. H. (2000). Functional anatomy of cognitive development: fMRI of verbal fluency in children and adults. *Neurology, 54,* 180–185. doi:10.1212/WNL.54.1.180

Gilbert, B., Belleville, S., Bherer, L., & Chouinard, S. (2005). Study of verbal working memory in patients with Parkinson's disease. *Neuropsychology, 19,* 106–114. doi:10.1037/0894-4105.19.1.106

Gordon, A. C. L., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68,* 70–83. doi:10.1006/jecp.1997.2423

Gregory, C., Lough, S., Stone, S., Erzinclioglu, S., Martin, L., Baron-Cohen, S., & Hodges, J. R. (2002). Theory of mind in patients with frontotemporal dementia and Alzheimer's disease: Theoretical and practical implications. *Brain, 125,* 752–764.

Henry, J. D., Phillips, L. H., Crawford, J. R., Ietswaart, M., & Summers, F. (2006). Theory of mind following traumatic brain injury: The role of emotion recognition and executive dysfunction. *Neuropsychologia, 44,* 1623–1628. doi:10.1016/j.neuropsychologia.2006.03.020

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

Kawamura, M., & Koyama, S. (2007). Social cognitive impairment in Parkinson's disease. *Journal of Neurology, 254*(Suppl. 4), IV49–IV53. doi:10.1007/s00415-007-4008-8

Knoke, D., Taylor, A. E., & Saint-Cyr, J. A. (1998). The differential effects of cueing on recall in Parkinson's disease and normal subjects. *Brain and Cognition, 38,* 261–274. doi:10.1006/brcg.1998.1042

Lees, A. J., & Smith, E. (1983). Cognitive deficits in the early stages of Parkinson's disease. *Brain, 106,* 257–270. doi:10.1093/brain/106.2.257

McNamara, P., Durso, R., Brown, A., & Lynch, A. (2003). Counterfactual cognitive deficit in persons with Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry, 74,* 1065–1070. doi:10.1136/jnnp.74.8.1065

Mengelberg, A., & Siegert, R. J. (2003). Is theory-of-mind impaired in Parkinson's disease? *Cognitive Neuropsychiatry, 8,* 191–209. doi:10.1080/13546800244000292

Mimura, M., Oeda, R., & Kawamura, M. (2006). Impaired decision making in Parkinson's disease. *Parkinsonism & Related Disorders, 12,* 169–175. doi:10.1016/j.parkreldis.2005.12.003

Monetta, L., Grindrod, C. M., & Pell, M. D. (2008). Effects of working memory capacity on inference generation during story comprehension in adults with Parkinson's disease. *Journal of Neurolinguistics, 21,* 400–417. doi:10.1016/j.jneuroling.2007.11.002

Monetta, L., Grindrod, C. M., & Pell, M. D. (2009). Irony comprehension and theory of mind deficits in patients with Parkinson's disease. *Cortex, 45,* 972–981. doi:10.1016/j.cortex.2009.02.021

Oeda, R. (2003). *Social cognition in Parkinson's disease* (Unpublished master's thesis). University of Tokyo, Tokyo, Japan.

Perón, J., Vicente, S., Leray, E., Drapier, S., Drapier, D., Cohen, R., . . . Verin, M. (2009). Are dopaminergic pathways involved in theory of

mind? A study in Parkinson's disease. *Neuropsychologia, 47,* 406–414. doi:10.1016/j.neuropsychologia.2008.09.008

Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors on false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development, 13,* 73–90. doi:10.1016/S0885-2014(98)90021-1

Sagar, H. J., Sullivan, E. V., Gabrieli, J. D. E., Corkin, S., & Growdon, J. H. (1988). Temporal ordering and short-term memory deficits in Parkinson's disease. *Brain, 111,* 525–539. doi:10.1093/brain/111.3.525

Saltzman, J., Strauss, E., Hunter, M., & Archibald, S. (2000). Theory of mind and executive functions in normal aging and Parkinson's disease. *Journal of the International Neurological Society, 6,* 781–788. doi:10.1017/S1355617700677056

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of a selective deficit in inhibiting self-perspective. *Brain, 128,* 1102–1111. doi:10.1093/brain/awh464

Shamay-Tsoory, S. G., & Aharon Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia, 45,* 3054–3067. doi:10.1016/j.neuropsychologia.2007.05.021

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Levkovitz, Y. (2007). The neuroanatomical basis of affective mentalizing in schizophrenia: Comparison of patients with schizophrenia and patients with localized prefrontal lesions. *Schizophrenia Research, 90,* 274–283. doi:10.1016/j.schres.2006.09.020

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10,* 640–656. doi:10.1162/089892998562942

Suzuki, A., Hoshino, T., Shigemasu, K., & Kawamura, M. (2006). Disgust specific impairment of facial expression recognition in Parkinson's disease. *Brain, 129,* 707–717. doi:10.1093/brain/awl011

Tsuruya, N., Kobayakawa, M., & Kawamura, M. (2011). Is "reading the mind in the eyes" impaired in Parkinson's disease? *Parkinsonism & Related Disorders, 17,* 246–248. doi:10.1016/j.parkreldis.2010.09.001

Watts, F. N., McLeod, A. K., & Morris, L. (1988). Associations between phenomenal and objective aspects of concentration problems in depressed subjects. *British Journal of Psychology, 79,* 241–250. doi:10.1111/j.2044-8295.1988.tb02285.x

Weintraub, D., Moberg, P. J., Culbertson, W. C., Duda, J. E., Katz, I. R., & Stern, M. B. (2005). Dimensions of executive function in Parkinson's disease. *Dementia and Geriatric Cognitive Disorders, 20,* 140–144. doi:10.1159/000087043

Werheid, K., Hoppe, C., Thone, A., Muller, U., Mungersdorf, M., & Yves von Cramon, D. (2002). The Adaptive Digit Ordering Test: Clinical application, reliability, and validity of a verbal working memory test. *Archives of Clinical Neuropsychology, 17,* 547–565.

Yoshimura, N., Kawamura, M., Masaoka, Y., & Homma, I. (2005). The amygdala of patients with Parkinson's disease is silent in response to fearful facial expressions. *Neuroscience, 131,* 523–534.

Zgaljardic, D. J., Borod, J. C., Foldi, N. S., Mattis, P. J., Gordon, M. F., Feigin, A., & Eidelberg, D. (2006). An examination of executive dysfunction associated with frontostriatal circuitry in Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology, 28,* 1127–1144. doi:10.1080/13803390500246910