CONFLICT AND FREEZING OF GAIT IN PARKINSON’S DISEASE: SUPPORT FOR A RESPONSE CONTROL DEFICIT

J. VANDENBOSSCHE,∗ a,b,c N. DEROOST, a,c E. SOETENS, a,c P. ZEISCHKA, a J. SPILDOOREN, d S. VERCRUYSSE, d A. NIEUWBOER d AND E. KERCKHOFS a,b,c

*Correspondence to: J. Vandenbossche, Cognitive Psychology, Vrije Universiteit Brussel, Brussels, Belgium

Abstract—We investigated response activation and suppression processes in Parkinson’s disease patients with freezing of gait (FOG). Fourteen freezers, 14 nonfreezers, and 14 matched healthy controls performed the attention network task (ANT) and the Stroop task. The ANT was designed to assess cognitive control, whereas the Stroop task was designed to assess automatic processes. Congruity effects were used as a general measure of conflict resolution. Supplementary reaction time (RT) distribution analyses were utilized to calculate conditional accuracy functions (CAFs) and delta plots to measure response activation and suppression processes. In agreement with previous research, freezes showed a general conflict resolution deficit compared with nonfreezers and healthy controls. Moreover, CAFs pointed to a strong initial incorrect response activation in FOG. As expected, conflict resolution impairment was only apparent in the ANT, and not in the Stroop task. These results suggest an imbalance between automatic and controlled processes in FOG, leading to a breakdown in both motor and cognitive response control. © 2012 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: response control, freezing of gait, Parkinson’s disease, conflict resolution, ANT, Stroop task.

Parkinson’s disease (PD) is associated with a number of prototypical cardinal motor features, such as tremor, rigidity, bradykinesia, and postural instability (Jankovic, 2008). In addition, major cognitive and affective disturbances are frequently reported in PD. Both motor and non-motor deficits result from a loss of dopaminergic cells in the basal ganglia, more specifically in the substantia nigra (Stocco et al., 2010). Five major neural circuits have been distinguished, which establish a connection between the basal ganglia and other projection areas. These pathways (motor, oculomotor, associative, limbic, and orbitofrontal), which are all disrupted in PD, account for many symptoms of the disease, as they are involved in a wide variety of functions including movement, attention, and learning (Obeso et al., 2008). A common movement problem, with both motor and cognitive components, is freezing of gait (FOG). FOG or motor blocks can be defined as a sudden and episodic inability to generate effective repetitive movements like stepping (Giladi and Nieuwboer, 2008). FOG contributes to the development of major disability and frequent falls, and can be considered as an independent cardinal sign of PD (Giladi et al., 2001). Giladi and Hausdorff (2006) described events that are prone to increase the occurrence of FOG episodes. According to these authors, these events can be motor based (advanced PD motor symptoms like a disordered step control), affective (depression, anxiety), or cognitive (dual tasking) in nature, stressing the wide range of triggers associated with FOG. Recently, Lewis and Barker (2009) proposed a pathophysiological model, trying to link FOG to motor, limbic, and cognitive brain loops. In this model, dopamine depletion results in an over-activation of the output nuclei in the substantia nigra, causing an inhibition on both the thalamus and pedunculopontine nucleus (PPN). This in turn impairs both ascending and descending pathways to brain areas involved in motor, cognitive, and limbic circuits. Increased limbic (stress or anxiety) and cognitive demands (working memory load) can lead to an over-activation of these circuits, resulting in sudden and intense episodes of excessive activation of the output nuclei on thalamus and PPN, triggering FOG. Moreover, the PPN and thalamus are also part of the cholinergic pathway, presuming to induce visuospatial and mnemonic deficits (Kehagia et al., 2010). So the involvement of the PPN could possibly be an important structure in understanding both the motor and cognitive components of FOG.

In a recent review, Nutt and colleagues (2011) postulated several hypotheses on the pathogenesis of FOG. One possible explanation is that the central drive and automaticity of movement is more affected in freezers compared with nonfreezers. This is also the key logic behind the current study: we expect freezers to be impaired in efficiently controlling automatic response activation, measured in congruency tasks, in comparison with nonfreezers and healthy controls. Evidence is growing that a variety of cognitive functions in PD patients with FOG are
deteriorated. Generalized executive dysfunction, deficits in cognitive flexibility, implicit learning impairment, and set-shifting difficulties under temporal pressure were found to be related to FOG (Amboni et al., 2008; Naismith et al., 2010; Vandenbossche, Deroost, Soetens, et al., unpublished observation). In a previous study (Vandenbossche et al., 2011), we already observed that FOG was associated with a general conflict resolution deficit. Fronto-striatal pathways (Beste et al., 2010) but also the neural pathways involving the PPN (Lewis and Barker, 2009) could lead to specific cognitive dysfunctions in PD. The current study aimed to elucidate on the process of conflict resolution by investigating response control in FOG patients. Response control refers to the set of processes that result in controlling incorrect reflex-like behavioral responses. This was studied by using two tasks, namely the Stroop task (Stroop, 1935; for a review see MacLeod, 1991) and the attention network test (ANT; Fan et al., 2002). In these so-called congruency tasks, irrelevant information interferes with relevant stimulus or response information. For example, in a Stroop task, participants have to respond to the color of the word (e.g. red, green, blue, or yellow) and to ignore its meaning (also red, green, blue, or yellow). The color and the meaning of the word can be identical (i.e. congruent) or different (i.e. incongruent). Generally, responses are faster and more accurate for congruent than for incongruent stimuli. A measure for conflict resolution is then obtained by subtracting reaction times (RTs) in the congruent condition from RTs in the incongruent condition. The ANT, on the other hand, is a combination of the cued reaction time task (Posner, 1980) and the flanker task (Eriksen and Eriksen, 1974). It has been developed to reliably test the efficiency of three attentional networks, namely alerting, orienting, and executive control (Posner and Rothbart, 2007). The executive-control network is a congruency measure representing conflict resolution. Although the attention networks are thought to be independent from each other (Fan et al., 2002), interactions with the executive-control network are sometimes observed (Callejas et al., 2004). Although the manual Stroop task has a number of potential sources of variance (e.g. perceptual conflict and semantic interference) compared with the ANT, a major difference is the level of stimulus–response (S–R) overlap (Kornblum et al., 1990). An irrelevant left- or rightward arrow, as used in the ANT, will elicit a strong automatic response activation, especially when responses are made by pressing a left and right key because of stimulus–response compatibility. The Stroop task, on the other hand, can be categorized as a stimulus–stimulus conflict task, with a strong overlap between the relevant stimulus (color of the word) and the irrelevant stimulus dimension (the word itself). Thus, in contrast with the ANT, automatic response activation will be much weaker in the manual Stroop task because the response set (four keys on the keyboard) does not share features with the relevant stimulus dimension (color of the word). Therefore, in the present study, both tasks are used to actively manipulate automatic response activation in PD patients and healthy controls.

The ANT and Stroop task can both be used to calculate congruency effects, reflecting conflict resolution. Congruency effects are generally explained by dual-processing models (e.g. de Jong et al., 1994; Eimer et al., 1995; Kornblum et al., 1999; Kornblum and Stevens, 2002; Ridderinkhof et al., 1995). In such models, relevant and irrelevant information are processed simultaneously along two routes, namely a controlled slower route based on the actual target stimulus, and an automatic faster route based on perceptual salience and stimulus–response associations. A response along the direct route will be activated earlier than a response along the deliberate route. In congruent task conditions, responses will be faster because there is no competition between routes that are both resulting in the correct response. For incongruent stimuli, however, responding through the direct fast route will lead to an error. These are the so-called fast errors, namely fast RTs, which are a reflection of unstoppable activation along the automatic route. Conditional accuracy functions (CAF), constructed by plotting accuracy rates against the RT distribution, allow detecting these fast errors and accordingly can be used to determine the strength of initial activation of the incorrect response (response activation). Some authors added a suppression or selective inhibition component to this dual-route architecture to further explain conflict resolution. For instance, Ridderinkhof (2002) formulated the activation–suppression hypothesis, which states that suppression processes are used to counter incorrect activation along the direct route, accumulating over time. With increasing response time, irrelevant response activation will increase, but also suppression itself will become more effective. As a consequence, the activation–suppression hypothesis predicts an increase in congruency effects with increasing response times that eventually should reduce, and in some instances, reverse at the slower segments of the distribution. Delta plots, in which congruency effects are plotted as a function of response speed, are useful to estimate suppression. Group differences in response activation (CAFs) and suppression (delta plots) for both the ANT and the Stroop task will thus provide more in-depth knowledge about response control in FOG.

It seems clear that congruency tasks provide a powerful context for examining conflict resolution during response and stimulus selection in PD. However, previous research assessing flanker effects in PD yielded mixed results. Several studies showed that both medicated and non-medicated PD patients demonstrated larger congruency effects compared with healthy controls (Praamstra et al., 1998, 1999; Wylie et al., 2005). In contrast, a number of other studies failed to find these apparent differences in interference control between PD patients and healthy controls (Cagigas et al., 2007; Falkenstein et al., 2006; Lee et al., 1999). A recent study by Wylie and colleagues (2009) investigated interference control in PD patients and healthy controls by means of distributional analyses. Departing from the activation–suppression hypothesis, they tried to provide new insights into the mechanisms of conflict resolution in PD. Results from this study revealed (a) stronger initial, incorrect response activation and (b) less...
efficient suppression in the activation of conflicting responses in PD. However, these effects were only observed in about half of the PD patients tested. In our view, FOG possibly be responsible for the observed variability in interference control deficits in PD (Vandenbossche et al., 2011). Given the above mentioned exaggerated executive deficits in freezers, it can be assumed that when this controlled route is hampered, freezers in particular, have to rely more on the automatic route. However, in case of an incongruent trial, more time is needed to overcome this strong response activation (slower RTs) resulting in greater congruency effects. Because of the possible imbalance between automatic and controlled processes in FOG, we expect that freezers are more susceptible to interference because they are hypothesized to experience stronger automatic activation of the incorrect response, leading to a response control deficit. This may also provide a partial explanation for why freezing of gait occurs in situations where response selection is required, such as when changing direction.

In the present study, we aimed to pinpoint a specific response control deficit in FOG. We expect to find larger general congruency effects for freezers compared with nonfreezers and healthy controls. Steeper CAFs determining increased capture of the response system by irrelevant information, and steadily increasing delta plots (i.e. positively going delta slopes) pointing to decreased efficiency of selective suppression in freezers compared with nonfreezers, could provide support for a response control deficit in FOG. If a stronger automatic response activation, or an impaired fine-tuning of the automatic–controlled route, is the source of congruency differences in FOG, we should observe more fast errors at the fastest segments of the RT distribution. Another possibility is that larger congruency effects are based on less efficient suppression of the automatic response as RTs get slower. In addition, we manipulated the amount of response control required to overcome conflict, by administering two types of congruency tasks (Stroop task and the ANT). The S–R overlap is stronger in the ANT compared with the Stroop task, putting more stress on the controlled processes attributed to overcome the incorrect response in case of an incongruent trial. If the hypothesis about an imbalance between automatic activation processes and controlled suppression is true, we expect that between-group differences and larger congruency effects would be more apparent in the ANT compared with the Stroop task because of differences in S–R overlap.

**EXPERIMENTAL PROCEDURES**

**Participants**

Twenty-eight PD patients (14 with FOG, freezers (FR); 14 without FOG, nonfreezers (nFR)) and 14 healthy controls (HC) participated in the study. All PD patients were diagnosed by a neurologist specialized in movement disorders. The new freezing of gait questionnaire (NFOGQ; Nieuwboer et al., 2009) score was constructed for three aims: (1) distinguishing freezers from nonfreezers, (2) rating the severity of freezing, and (3) rating the impact of FOG on daily life. A video, showing several examples of FOG, was presented to increase the recognition of this phenomenon in patients. When patients experienced FOG at least once during the last month, they scored above zero on the NFOGQ and were accordingly assigned to the FR group. All participants had normal to corrected-to-normal vision, did not have a color vision deficiency, had no deep brain stimulation, and had no orthopedic or additional neurological disorders. Participation in the experiment was in accordance with the Ethics Committee of the Vrije Universiteit Brussel (VUB).

All three groups (FR, nFR, HC) were matched for age, gender, and education (see Table 1). All participants scored above the standard cutoff score of 24 on the mini mental state examination (MMSE; Folstein et al., 1975), an indication that overt dementia symptoms were absent. In addition, the scale for outcomes in Parkinson’s disease-cognition (SCOPA-COG; Marinus et al., 2003) was used to assess patients’ global cognitive capacities. The SCOPA-COG contains 10 items, including subscales such as “memory” (replicating the order in which cubes were pointed out, digit span backward, immediate, and delayed word recall), “attention” (counting down by threes and months backward), “executive functioning” (successive repetitions of fist-edge-palm movements, set shifting with dice and fluency animals), and “visuospatial functioning” (mental reconstruction of figures). Low scores on these subtests are indicative of a deficit. Finally, a measure of affective disturbance was obtained by administering the hospital anxiety depression scales (HADS; Zigmond and Snaith, 1983). Scores higher than seven points on HADS subscales indicated increased complaints associated with anxiety and depression.

**Material and procedure**

The study was conducted in participants’ home setting under supervision of the experimenter. The congruency experiments (ANT and Stroop task) were run on an Intel core 2 duo portable computer with 15.6 inch screen, using E-prime version 1.1 software (Schneider et al., 2002), after signing of informed consent. All participants were asked to complete the ANT and the Stroop task in one session, with a pause of 15 min between both experiments. Testing occurred solely in ON-phase, about 60-90 min after patients took their morning dose of anti-Parkinson medication.

**ANT.** Participants were seated approximately 65 cm from the computer screen. They were instructed to focus on a centrally located fixation cross throughout the task and to respond as fast and as accurately as possible to the left–right direction of a central black arrow (target). Flanker stimuli surrounding the target consisted of a row of four black horizontal arrows pointing leftward or rightward, which were presented in a location above or below the fixation point against a white background. The “c” and “n” keys, situated on the bottom row of an AZERTY keyboard, corresponded to the left and right direction of the target and had to be pressed with the left and right index finger, respectively. The four identical flanker arrows (two on each side of the target), could point in the same (congruent condition) or opposite (incongruent condition) direction as the central target arrow. In the neutral condition the flankers were four lines without arrowheads.

The total duration of a trial lasted 4000 ms and consisted of five events (see Fig. 1). First, a fixation cross was presented in the center of the screen for a random variable fixation period between 400 and 1600 ms. Then a warning cue (asterisk) appeared for 100 ms. There were four types of warning cues (see Fig. 1): no cue, center cue (participants were shown an asterisk at the location of the fixation cross), double cue (two asterisks appeared; one above and one below fixation cross), or a spatial cue (an asterisk was presented at the location where the target would appear). After the warning cue, there was a short fixation period of 400 ms followed by the stimulus display, consisting of the five arrows (four flankers and one central target), presented above or below the fixation...
cross. The stimulus was presented until the participant responded, but no longer than 1700 ms. After a response was made, target and flankers disappeared immediately. The next trial began after a variable post target fixation period, based on the duration of the variable fixation period at the beginning of the trial and the RT.

Before starting the experimental ANT blocks, participants completed a practice block of 24 trials in a random order. After practice, they executed three experimental blocks, each consisting of 96 trials (four cue conditions/2 target locations/2 target directions/3 flanker conditions/2 repetitions). The presentation of trials in each condition was random. RTs and accuracy were recorded in each trial. Feedback on accuracy was given during the practice block only. Participants were not warned when they made an error and received no feedback about their error rates and RTs. The practice block took about 2 min and the experimental blocks about 18 min to accomplish. Participants were given the possibility to rest between blocks.

**Stroop task.** Participants were seated approximately 65 cm from the computer screen. A trial started with the presentation of a word stimulus in the center of the screen. A word stimulus could randomly consist of one of four color words (red, green, blue, or yellow), and was displayed in one of four colors (red, green, blue, or yellow). Words were depicted in font Courier New, font size 18,

Table 1. Clinical, motor, and neuropsychological measurements across PD groups (FR and nFR) and healthy controls (HC)

<table>
<thead>
<tr>
<th>Measure</th>
<th>FR (n=14)</th>
<th>nFR (n=14)</th>
<th>HC (n=14)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (M:F)</td>
<td>11/3 (79% male)</td>
<td>11/3 (79% male)</td>
<td>11/3 (79% male)</td>
<td></td>
</tr>
<tr>
<td>H&amp;Y (OFF) (2; 2.5; 3)</td>
<td>(36%; 36%; 28%)</td>
<td>(50%; 43%; 7%)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Levodopa therapy (n)</td>
<td>14 (100%)</td>
<td>13 (93%)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Adjunct therapy (n)</td>
<td>12 (86%)</td>
<td>13 (93%)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Disease duration (y)</td>
<td>10.79 0.85</td>
<td>8.21 0.91</td>
<td>NA .048*</td>
<td></td>
</tr>
<tr>
<td>UPDRS–III (OFF)</td>
<td>36.36 3.52</td>
<td>33.21 2.07</td>
<td>NA .448</td>
<td></td>
</tr>
<tr>
<td>Age (y)</td>
<td>65.72 2.11</td>
<td>68.03 1.37</td>
<td>66.30 1.86</td>
<td>.701</td>
</tr>
<tr>
<td>Education (years in school)</td>
<td>19.07 0.85</td>
<td>20.21 0.99</td>
<td>20.14 0.48</td>
<td>.534</td>
</tr>
<tr>
<td>Mini Mental State Examination (ON)</td>
<td>28.00 0.28</td>
<td>28.79 0.37</td>
<td>29.29 0.30</td>
<td>.023*</td>
</tr>
<tr>
<td>HADS—anxiety (ON)</td>
<td>7.00 0.90</td>
<td>5.64 0.95</td>
<td>5.00 1.32</td>
<td>.413</td>
</tr>
<tr>
<td>HADS—depression (ON)</td>
<td>5.86 1.06</td>
<td>6.07 0.74</td>
<td>4.14 1.05</td>
<td>.312</td>
</tr>
<tr>
<td>SCOPA-COG (ON)</td>
<td>27.56 1.33</td>
<td>30.43 1.15</td>
<td>32.79 1.31</td>
<td>.031*</td>
</tr>
<tr>
<td>(1) Memory and learning</td>
<td>10.50 0.86</td>
<td>11.71 0.84</td>
<td>13.43 1.07</td>
<td>.096</td>
</tr>
<tr>
<td>(2) Attention</td>
<td>3.14 0.21</td>
<td>3.71 0.16</td>
<td>4.00 0.00</td>
<td>.001**</td>
</tr>
<tr>
<td>(3) Executive functions</td>
<td>9.71 0.57</td>
<td>10.64 0.31</td>
<td>10.86 0.38</td>
<td>.152</td>
</tr>
<tr>
<td>(4) Visuospatial functions</td>
<td>4.50 0.14</td>
<td>4.36 0.17</td>
<td>4.50 0.17</td>
<td>.771</td>
</tr>
</tbody>
</table>

M, mean; SE, standard error of mean; NA, not applicable; H&Y, Hoehn and Yahr rating scale; UPDRS–III, unified Parkinson’s disease rating scale; HADS, hospital anxiety depression scales; SCOPA-COG, Scales for Outcomes in Parkinson’s disease-COGnition.

* Significant at 0.05 level; ** Significant at 0.001 level.

Fig. 1. The Attention Network Test (ANT), adapted from Fan and colleagues (2002). The experimental procedure is explained with an example of an incongruent trial, the four possible cues, and the three flanker conditions. The calculation of the three ANT networks is shown in the bottom right corner of the figure.
and were presented individually against a white background. The color word was presented until a response was given within the allowed response window of 5000 ms. Participants were instructed to react as fast and as accurately as possible to the color identity of the word and to ignore its meaning. RTs and accuracy were recorded on each trial. Participants had to press the “c,” “v,” “b,” and “n” keys with the left middle finger, left index finger, right index finger, and right middle finger for a red, green, blue, and yellow color, respectively. Participants were explicitly instructed not to read the word presented. In case of an incorrect response, the word “Error” was presented in Dutch for 760 ms. No error corrections were possible. The next color word appeared after a response–stimulus interval of 200 ms. In 75% of the trials, color and meaning of the word did not match (incongruent condition); in the remaining 25% of the trials color and meaning of the word matched (congruent condition). Participants first completed two practice blocks of 50 trials, followed by five experimental blocks of 100 trials. After each block of trials, patients received feedback about their error rates and RTs. A break of 30 s was imposed before the next block started.

Statistical analysis

Independent t-tests, one-way analysis of variances (ANOVAs), and a correlation analysis (Pearson and Spearman rho statistics) were used to analyze group differences for clinical, motor, and demographical measures and its relationship with severity of freezing. Mixed factorial ANOVAs (with Huynh–Feldt corrections for violations of sphericity) were implemented to analyze ANT and Stroop performance. Bonferroni post hoc tests were performed in case of significant group differences. Accuracy rates for the ANT and Stroop task were arcsine transformed. Data analysis was performed using SPSS version 17.0, and all analyses were two-tailed, using a significance level of .05. Differences showing a significance level of .10 will be reported as a tendency.

RESULTS

Clinical, motor, and demographical differences

At the time of testing, all patients were in the ON-phase of anti-Parkinson medication. Medications used by patients consisted mostly of levodopa products and adjunct therapies. One nonfreezer and one freezer were on anti-cholinergic treatment, and two patients (one in each group) were taking a selective serotonin reuptake inhibitor.

We compared motor symptoms between both PD groups by means of the Hoehn and Yahr (1967) scale, providing a gross assessment of disease progression through several stages (0–5) going from no signs of the disease to complete dependency, and Section III of the unified Parkinson’s disease rating scale (UPDRS; Fahn et al., 1987). These tests were assessed in the OFF phase when the action of medication was strongly decreased or absent. No apparent difference could be observed between groups. However, freezers differed significantly in disease duration compared to nonfreezers (FR > nFR). Severity of freezing (FR: mean (M) = 13.2; standard error of mean (SE) = 1.95), investigated by means of the NFOGQ, revealed a positive correlation with disease duration (r = .56, p < .05), Hoehn and Yahr scale (r = .60, p < .05), HADS anxiety (r = .69, p < .05), and HADS depression (r = .71, p < .01).

Regarding the cognitive status of participants, we observed that both the MMSE and SCOPA-COG yielded significant differences between groups. Post hoc Bonferroni tests showed that MMSE and overall SCOPA-COG scores were higher for healthy controls as compared with freezers (P < .05) Performance on the subtest of the SCOPA-COG covering attention differed between freezers and both nonfreezers and controls (FR < nFR and HC). However, correlations between SCOPA-COG scores and congruency measures were not significant. Dissimilarities between groups for the remaining general descriptive (including medication profiles) and cognitive measures failed to reach significance, indicating comparable scores for all participants.

ANT

Median RTs of correct trials in each condition (no cue, double cue, center cue, spatial cue, neutral flanker, incongruent flanker, and congruent flanker) were used to assess the three attentional networks (alerting, orienting, and executive control). Erroneous responses and responses following an error were discarded from the analysis. The executive-control network is assumed to represent a measure for cognitive control, as participants need to control conflict by inhibiting an irrelevant response. Alerting and orienting networks represent the ability to maintain a vigilant state and focus attention. The ANT provides the opportunity to dissociate these three attentional networks on the basis of RT differences. A larger difference between two conditions suggests more impairment of the corresponding network.

Error rate analysis showed that overall accuracy rates did not differ significantly between groups (FR: M = 97.10%, SE = .69; nFR: M = 97.10%, SE = 1.02; HC: M = 98.34%, SE = .30; F(2,41) = .85, P = .44). Moreover, FRs did not make more omission errors under conflict (i.e. under incongruent trials) as compared with nFRs and HCs, F(2,41) = .43, P = .66. Therefore, network effects were determined on the basis of the RT analysis only.

Attentional networks. For the executive control network, we applied a 3 × 2 repeated measures ANOVA with group as between-subject factor (FR, nFR, and HC), and flanker type (congruent and incongruent flanker trials) as within-subject factor. The main group effect was not significant, indicating that the RT level did not differ between groups, F(2,39) = 1.03, P = .37, ηp2 = 0.05. A main effect for the executive-control network was found, meaning that responses were faster in congruent as compared with incongruent flanker trials, F(1,39) = 32.54, P < .001, ηp2 = .49. More importantly, we found a significant group × executive-control network interaction F(2,39) = 5.51, P < .01, ηp2 = .22, revealing that congruency differences between groups were present (see Fig. 2A). A Bonferroni post hoc test confirmed that freezers showed a greater RT difference between congruent and incongruent flanker trials (M = 130.29 ms, SE = 10.72) compared with nonfreezers (M = 86.36 ms, SE = 6.84; P = .01) and controls (M = 95.43 ms, SE = 11.45; P = .05). Congruency differences between nonfreezers and controls were not significant, P = 1.
A possible concern is that executive control and disease duration are confounded \((r=.48, P=.01)\). Therefore, we carried out a \(2 \times 2\) repeated measures ANCOVA with group as between-subject factor (FR and nFR), flanker type (congruent and incongruent flanker trials) as within-subject factor and disease duration as covariate. Similar to our ANOVA results we found no significant main effect of group \((F(1,25)=0.24, P=.88, \eta^2=.001)\), a main effect for the executive-control network \((F(1,25)=15.01, P=.001, \eta^2=.38)\). Also the group×executive-control network interaction \((F(1,25)=7.02, P<.05, \eta^2=.22)\) remained significant. These results indicate that even after controlling for disease duration, differences between freezers and nonfreezers on the executive-control network persist.

For the alerting and orienting network, we performed two separate \(3 \times 2\) repeated measures ANOVAs with group as between-subject factor (FR, nFR, and HC), and cue type (alerting: no- and double-cue trials; orienting: center- and spatial-cue trials) as within-subject factor. The main group effects were not significant, indicating that RT levels did not differ between groups (alerting: \(F(2,39)=.97, P=.39, \eta^2=.05\); orienting: \(F(2,39)=1.21, P=.31, \eta^2=.06\)). Main effects for the alerting and orienting network were significant, indicating that responses were faster in double- as compared with no-cue trials (alerting: \(F(1,39)=31.14, P<.001, \eta^2=.49\) and responses were faster in spatial- as compared with center-cue trials (orienting: \(F(1,39)=121.95, P<.001, \eta^2=.76\)). Group×alerting and group×orienting interactions were not significant, respectively \(F(2,39)=0.13, P=.88, \eta^2=.01\) and \(F(2,39)=0.77, P=.47, \eta^2=.04\). These statistics indicated that alerting scores (FR: \(M=30.02\) ms, SE=5.96; nFR: \(M=35.59\) ms, SE=5.71; HC: \(M=29.86\) ms, SE=5.18) and orienting scores (FR: \(M=41.65\) ms, SE=9.04; nFR: \(M=45.39\) ms, SE=7.44; HC: \(M=51.46\) ms, SE=9.82) did not differ between groups.

**Interactions with the executive control network.** As the ANT is a complex task designed to measure more than one aspect of attention, interactions between attentional networks can be observed \((Callejas et al., 2004)\). Therefore it is important to check whether the executive-control network is modulated by alerting or orienting cues. We analyzed these interactions by means of two separate \(3 \times 2\) repeated measures ANOVAs with group as between-subject factor (FR, nFR, and HC), and the congruency effect for a specific cue type (alerting: congruency effect for no- and double-cue trials; orienting: congruency effect for center- and spatial-cue trials) as within-subject factor.

For the alerting and executive-control network interaction we found a tendency toward a larger congruency difference between congruent and incongruent flanker trials when an alerting cue was present (double-cue trials) compared with trials when this was not the case (no-cue trials), \(F(1,39)=3.68, P=.06, \eta^2=.09\). More importantly, no group differences for the alerting×executive-control network interaction could be found, \(F(2,39)=0.01, P=.99, \eta^2<.01\).

For the orienting- and executive-control network interaction we could not discover a significant difference between congruent and incongruent flanker trials when an orienting cue was offered (spatial-cue trials) compared with uncued trials (center-cue trials), \(F(1,39)=0.46, P=.50, \eta^2=.01\). Moreover, group differences for the orienting×executive-control network interaction were not significant, \(F(2,39)=0.86, P=.43, \eta^2=.04\).

**Stroop task**

An RT analysis was performed on participants’ median RTs in incongruent and congruent conditions, with the exclusion of practice trials, to assess the Stroop effect. Erroneous responses and responses following an error were discarded from the analysis. Two freezers were left out of the analysis because they were unable to complete the task because of motor limitations. Therefore we only included 12 patients in our FR group, as opposed to 14 nonfreezers and 14 healthy controls. A larger Stroop effect suggests decreased cognitive control.

Error rate analysis shows that overall accuracy rates differed significantly between groups (FR: \(M=92.61\%), SE=1.28; nFR: M=94.07\%, SE=1.17; HC: M=96.84\%, SE=2.54; \(F(2,39)=4.99, P<0.05\)). A post hoc Bonferroni test indicated that freezers were making significantly more errors than healthy controls, \(P<.05\). When we analyzed incongruent and congruent trials separately, we observed that differences between groups remained. A post hoc Bonferroni showed that freezers made more errors in congruent trials compared with healthy controls \((P<.05)\), FR: \(M=93.74\%, SE=1.05; nFR: M=96.51\%, SE=0.58; HC:
M=97.29%, SE=0.75; F(2,39)=4.82, P<.05. For incongruent trials, we can only observe that healthy controls tended to make less errors compared with nonfreezers (P=.08) and freezers (P<.05) (FR: M=91.47%, SE=1.67; nFR: M=91.63%, SE=1.97; HC: M=96.38%, SE=0.73; F(2,39)=4.01, P<.05). Moreover, freezers did not make more omission errors under conflict (i.e. under incongruent trials) compared with nonfreezers and healthy controls, F(2,39)=1.26, P=.30. Because of the similarity in patterns between RTs and errors, Stroop effects will be derived from the RT analysis only.

We used a 3×2 repeated measures ANOVA with group as between-subject factor (FR, nFR, and HC), and congruency (congruent and incongruent trials) as within-subject factor. We did not observe a significant group congruency interaction F(2,37)=0.97, P<.39, η²=0.05, meaning that there were no differences in Stroop effects between groups (FR: M=200.17 ms, SE=30.66; nFR: M=230.64 ms, SE=38.51; HC: M=170.79 ms, SE=21.66). The main group effect only showed a tendency toward significance, indicating that the RT level tended to differ between groups, F(2,37)=2.59, P=.09, η²=0.12. A main congruency effect was found, meaning that responses were slower in incongruent compared with congruent color trials, F(1,37)=123.53, P<.001, η²=0.77 (see Fig. 2B).

**Activation–suppression hypothesis: distributional analyses**

Regarding the activation-suppression hypothesis, we expected that participants who experienced strong automatic response activation would make more fast errors at the fastest bins of the RT distribution for the incongruent condition. We captured the differences in fast errors between groups by measuring the strength of the initial response activation. For each group, accuracy rates for the two fastest bins in the RT distribution were calculated (CAFs). A steeper slope between Bin 1 and Bin 2 reflected stronger incorrect response activation.

To assess the efficiency of suppressing an activation of an incorrect response, a process that comes into play after surpassing a certain activation threshold, we analyzed the slowest bins of the RT distribution. We expected that groups with less efficient suppression would show an increase in congruency effects throughout the whole RT distribution (i.e. positively going delta slopes). In other words, they lack the ability to suppress the incorrect response activation, as their RTs become slower. We calculated group differences in suppression by comparing congruency differences for the slowest bins in the RT distribution (delta plots).

To calculate CAFs and delta plots, we studied RT distributions and divided our data into six RT bins and then averaged across subjects to obtain group sextiles (de Jong et al., 1994; Ridderinkhof, 2002).

**Response activation.** ANT. Conditional accuracy functions were calculated for incongruent trials and marked to analyze the pattern of fast errors. This was done by rank-ordering RTs for all responses to an incongruent trial and dividing them into six equal size bins. Accuracy rates were then calculated separately for each bin and plotted against the mean RT for each bin (see Fig. 3A).

We conducted a 3×2 repeated measures ANOVA with group as between-subject factor (FR, nFR, and HC), and bin (bin 1 and bin 2) as within-subject factor. The main group effect was not significant, indicating that the general accuracy rate did not differ between groups, F(2,39)=0.28, P=.76, η²=0.01. A main effect for bin was found, meaning that responses were less accurate in the fastest (Bin 1) compared with Bin 2, F(1,39)=38.23, P<.001, η²=0.50. More importantly, we also found a significant group×bin interaction, F(2,39)=3.35, P<.05, η²=0.15, meaning that differences in fast errors between groups were present. A Bonferroni post hoc test revealed that freezers showed more fast errors (M=14.65%, SE=2.96) as compared with healthy controls (M=4.03%, SE=3.04; P<.05), whereas nonfreezers did not differ significantly from either group (M=9.14%, SE=3.23; P>.10).

Stroop task. Conditional accuracy functions were calculated in the same manner as for the ANT data. Accuracy rates were computed separately for each bin and plotted against the mean RT for each of the six bins (see Fig. 3B).

Graphically, we noticed that for the PD groups most errors were made at the sextiles containing the shortest and the longest response latencies. Healthy controls, however, expressed a somewhat surprising pattern because a high accuracy rate was observed in the fastest bin and...
more errors were made for the slowest RTs. We conducted a 3×2 repeated measures ANOVA with group as between-subject factor (FR, nFR, and HC), and bin (bin 1 vs. bin 2) as within-subject factor. The main group effect was significant, indicating that general accuracy rate level differed between groups, F(2,37)=3.16, P=.05, ηp²=0.15. However, neither main bin effect nor a significant group×bin interaction was found, indicating that responses were equally accurate in the fastest bin compared with the next bin for all groups, F(1,37)=0.27, P=.60, ηp²<0.01 and F(2,37)=1.78, P=.18, ηp²=0.09, respectively.

For the Stroop task, we also determined response activation between groups by comparing average accuracy rates for the first bin. A one-way ANOVA with group (FR, nFR, and HC) as between-subject factor showed that the average accuracy rate in the first bin differed significantly between groups, F(2,39)=6.11, P<0.01. A post hoc Bonferroni test demonstrated that healthy controls (M=98.12%, SE=0.73) made less errors in the first bin as compared with freezers (M=91.18%, SE=2.79; P<.05) and nonfreezers (M=90.86%, SE=2.96; P<.05). The FR and nFR groups, however, did not differ significantly, P=1.

Following these results, differences regarding automatic response activation between groups are less apparent in the Stroop task compared with the ANT.

Response suppression. ANT. RT delta plots were constructed to investigate the efficiency of suppression counteracting interference between incongruent and congruent stimuli. These plots denote congruency effects as a function of mean RT. To calculate rank-ordered RTs for all responses to congruent and incongruent trials, we divided them into six equal-size bins. Congruency effects (delta values) were calculated separately for each bin by subtracting the mean RT for the congruent condition from the mean RT for the incongruent condition and then plotting the difference (delta) against the mean RT for each bin (see Fig. 4A).

Graphically, we observed that nonfreezers and controls demonstrated a similar pattern of steady increase in congruency effect throughout the whole RT distribution, as opposed to freezers who show a gentle decrease toward the last bin. We tested possible differences between groups by running a 3×2 repeated measures ANOVA with group as between-subject factor (FR, nFR, and HC), and bin (bin 5 and bin 6) as within-subject factor. We found no significant main or interaction effects (main effect group: F(2,39)=1.59, P=.22, ηp²=0.08; main effect bin: F(1,39)=0.33, P=.57, ηp²<0.01; interaction group×bin: F(2,39)=0.82, P=.45, ηp²=0.04). To test whether the decrease in congruency effect for the last bin in freezers was significantly different from the pattern expressed by nonfreezers and healthy controls, we additionally conducted a 3×2 repeated measures ANOVA with group as between-subject factor (FR, nFR, and HC), and bin (delta slope bin 5–4 and delta slope bin 6–5) as within-subject factor. However, main or interaction effects remained insignificant, indicating an absence of response suppression in all three groups (main effect group: F(2,39)=0.82, P=.45, ηp²=0.04; main effect bin: F(1,39)=0.41, P=.53, ηp²=0.01; interaction group×bin: F(2,39)=0.67, P=.52, ηp²=0.03).

Stroop task. RT delta plots were constructed in the same manner as for the ANT data. Congruency effects (delta values) were calculated separately for each of the six bins by subtracting mean RT for the congruent condition from mean RT for the incongruent condition, which were then plotted against the mean RT for each bin (see Fig. 4B).

Graphically, we see that all groups demonstrated a similar pattern of steady increase of the congruency effect across the fastest and intermediate parts of the RT distribution followed by a decrease toward the last bin. A 3×2 repeated measures ANOVA with group as between-subject factor (FR, nFR, and HC), and bin (bin 5 and bin 6) as within-subject factor, however, revealed no significant main or interaction effects (main effect group: F(2,37)=1.04, P=.36, ηp²=0.05; main effect bin: F(1,37)=0.30, P=.59, ηp²<0.01; interaction group×bin: F(2,37)=0.03, P=.97, ηp²<0.01). To investigate the decrease in the last bin, we conducted a 3×2 repeated measures ANOVA with group as between-subject factor (FR, nFR, and HC), and bin (delta slope bin 5–4 and delta slope bin 6–5) as within-subject factor. There was no significant main group effect, meaning that the delta slope level was comparable between groups, F(2,37)=0.07, P=.94, ηp²<0.01. Importantly, there was a significant main effect of bin showing a significant decrease in slope between bins 5–4 and bins 6–5, F(1,37)=4.75, P<.05, ηp²<0.11. The group×response suppression interaction, however, did not reach significance, meaning there were no group differences in response sup-
expression, \( F(2,37)=0.05, P=.95, \eta_p^2=0.03 \). These results provide evidence for equal effective suppression for all groups in the Stroop task.

**GENERAL DISCUSSION**

In the present study we investigated conflict resolution in FOG by means of the ANT and the Stroop task. Differences in S–R overlap between both tasks were assumed to yield differences in automatic response activation strength, requiring more (ANT) or less (Stroop task) control. RT distribution analyses were used to clarify conflict resolution and disentangle the effects of automatic response activation and suppression processes. As expected, larger general congruency effects for freezers compared with nonfreezers and healthy controls were apparent for the executive-control network in the ANT. This is in agreement with our previous findings (Vandenbossche et al., 2011), and supports the idea of impaired conflict resolution associated with FOG. When S–R overlap is low (Stroop task), response control is less stressed, leading to no significant group differences in general congruency effects. Distributional analyses, carried out to assess the activation–suppression hypothesis (Riddoch, 2002), also showed discrepancies between the ANT and Stroop task. For one, support for a stronger automatic response activation in FOG was only found in the ANT. On the other hand, difference in an increased suppression of the activated response for the slowest RTs was not observed in freezers. This is rather surprising because suppression is expected to be found when S–R overlap is larger, for example in a Simon or flanker task. As we used a within-subjects design, where all subjects performed both the ANT and the Stroop task, individual differences could not explain these mixed findings. Our results suggest that freezers are more led by automatic processes when controlled processes are stressed.

The ANT and the Stroop task were administered to measure congruency effects in PD patients and healthy controls. For Stroop performance, a significant difference between congruent and incongruent trials was found for all groups, albeit, not differing between groups. Important to mention is that overall, freezers made more errors as compared with healthy controls, whereas nonfreezers were situated in between. This accuracy rate difference between groups was not present in the ANT, possibly indicating that freezers had more difficulty with the Stroop task than with the ANT. For ANT performance, we demonstrated that freezers were more impaired at the executive-control network compared with both nonfreezers and healthy controls. This is in agreement with our previous research (Vandenbossche et al., 2011) reflecting a general conflict resolution deficit in FOG. Although freezers showed a decreased score on the attention subscale of the SCOPA-COG, this deficit was not reflected in the alerting and orienting network scores in the ANT. Previous research showed specific interactions between the attentional networks (Callejas et al., 2004). In our study, only a tendency toward an interaction between the alerting- and executive-control network was found in the sense of a larger congruency difference when an alerting cue (double cue) was presented compared with trials in which it was absent. However, as there were no differences in interactions between groups, group differences in the executive-control network can be considered as genuine.

Distributional analyses provided insight into general conflict resolution differences between groups. ANT data showed that patients expressing freezing of gait are relying less on the controlled deliberate route when facing conflict. This is probably because the deliberate controlled route, modulated by the striatum (Poldrack et al., 1999; Van der Graaf et al., 2004) is hampered in PD patients, and freezers in particular. Given the specific cognitive problems in FOG concerning executive functioning and controlled processing (Amboni et al., 2008; Naismith et al., 2010), freezers experienced stronger automatic response activation, leading to more fast errors and larger congruency effects. When less S–R overlap was present, as in the Stroop task, no differences in automatic response activation occurred between groups. Stronger initial, incorrect response activation could not be coupled with a reduced ability to suppress this activation (steeper delta slopes). These results are partly in accordance with the results of Wylie and colleagues (2009), not only showing a stronger initial response activation in PD, but also impaired suppression. The suppression mechanism, however, has to be interpreted with caution. Flat delta slopes, for example, can indeed reflect more suppression, but also a diminished increase in irrelevant response activation (Zeischka et al., 2011). Moreover, Wylie and colleagues found these deficits for about half of the PD patients tested. As freezers clearly show a similar pattern of impairment, FOG may account for the observed differences in PD patients in their study.

General screening tests show that freezers scored lower than healthy controls on the MMSE. However, all freezers scored above the cutoff of 24, indicating that no one suffered from dementia. Overall SCOPA-COG and the score on the attention subscale also differed between FRs and HCs, but all group results fall within the normative data collected in the PROPARK study (Verbaan et al., 2007), including the FOG patients. Also no correlations between SCOPA-COG scores and congruency measures were present. Although participants were matched for age, gender, and education, we noticed that freezers have on average been longer diagnosed with PD as compared with nonfreezers. However, UPDRS-III scores, the best representation of disease status, support the notion that freezers and nonfreezers had comparable disease profiles. Correlation analyses revealed that severity of freezing, as measured with the NFOGQ, was associated with disease duration, the Hoehn and Yahr scale and scores on the HADS (anxiety and depression). This fits well with the conceptual framework proposed by Giladi and Hausdorff (2006), linking motor, cognitive, and affective aspects to FOG.

Although the present study indicates clearly that freezers rely more on automatic response activation instead of...
on the deliberate controlled route, it might be interesting to investigate whether FOG patients only experience stronger initial motor activation, or that they just are less able to control an incongruent task condition. A possible way to investigate this, is by using EEG measurements. Ample research studying response conflict through event-related potentials showed specific deficits in PD (for example Bokura et al., 2005). However, EEG measurements in PD patients expressing FOG are rarely explored. We propose to investigate lateralized readiness potentials (LRPs), reflecting the preparation of motor activity, to further disentangle activation and suppression processes in conflict resolution for patients expressing FOG. Another important objective could be the understanding of the role of the PPN in the specific cognitive deficits in FOG. As already mentioned, these structures could play a key role in motor and cognitive features of FOG (Lewis and Barker, 2009; Kehagia et al., 2010). Future studies should try to clarify which brain areas and/or structures can be linked with response control deficits in FOG using brain imaging data identifying the (in)activity of alternative brain regions while freezers are confronted with response conflict.

A possible issue is that in the current study, task administration (ANT and Stroop task) was not counterbalanced. We acknowledge that it is important to exclude possible ANT contamination in the Stroop task. However, MacLeod (1991) concluded that Stroop training effects are mixed. For example, subjects were offered Stroop task training for 8 days: initially, interference was very strong, then it reduced quickly over the first couple of days, and finally it declined more gradually. Overall, interference from incompatible trials was still present after 8 days of training. Moreover, MacLeod (1998) stated that Stroop training effects are highly specific and do not generalize beyond the particular stimuli trained. Therefore, it seems unlikely that ANT performance would affect Stroop performance in the current study. First, differences in the amount of S–R overlap and the nature of both tasks (arrows in the ANT, words and colors in the Stroop task) would hamper the generalization of practice. Second, the response set is also very different: the ANT is a two-choice RT task and the manual Stroop task is a four-choice RT task. Third, the ANT only lasted for 30 min, which is not comparable with several days of training needed to find an effect within the Stroop task (MacLeod, 1991). Finally, both tasks were separated with a break going from 30 to 60 min. Although on the basis of previous research we would not expect an influence, we suggest that future studies should counterbalance task order to control for any effects, as subtle as they may be. All participants first executed the ANT task and afterward began the Stroop task. Although the response set is very different between both tasks training could still occur. Indeed, without counterbalancing, we cannot totally exclude that ANT training effects have been carried over to the Stroop task. However, training effects in Stroop tasks are very inconsistent and have not yet been found to generalize beyond the particular stimuli trained (MacLeod, 1991, 1998).

In conclusion, we demonstrated that freezers are less efficient in holding an equilibrium between automaticity and controlled processing, important in resolving conflict. Freezers are thought to rely more on automatic processing because executive functions are more impaired. When conflict does not elicit a strong response activation, like in the Stroop task, freezers are able to control interference in an adequate way. However, as soon as actions become more ambiguous and complex, automatic and controlled processes seem to work less synchronized. This in turn leads to a breakdown in both motor action and cognitive processing. The freezing episode can thus be seen as a kind of involuntary time-out, urging freezers to get back on the right track.

Acknowledgments—We thank Prof. Dr. Wim Vandenberghe for his assistance in recruiting PD patients. This research was funded by the Flanders Research Funds (G.0691.08) and the Research Council of the Vrije Universiteit Brussel (grant OZR-1933 BOF).

REFERENCES


