THE EFFECTS OF LEE SILVERMAN VOICE TREATMENT ON THE FACIAL MOVEMENT OF PARKINSON'S DISEASE PATIENTS AND THE WAY THEY ARE PERCEIVED BY OTHERS

by

ALEKSEY I. DUMER

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2011

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Joan C. Borod, Ph.D.

Date — Chair of Examining Committee

Maureen O'Connor, Ph.D., J.D.

Date — Executive Officer

Joan C. Borod, Ph.D.

Harriet Oster, Ph.D.

Laura Rabin, Ph.D.

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

THE EFFECTS OF THE LEE SILVERMAN VOICE TREATMENT ON THE FACIAL MOVEMENT OF PARKINSON'S DISEASE PATIENTS AND THE WAY THEY ARE PERCEIVED BY OTHERS

by

Aleksey I. Dumer

Adviser: Professor Joan C. Borod, Ph.D.

Parkinson's disease (PD) is characterized by impaired facial movement, a deficit that, as previous studies suggest, leads others to attribute negative traits to PD patients. Given the associations between facial movement and vocal parameters, it was hypothesized that the Lee Silverman Voice Treatment (LSVT; Ramig, Fox, & Sapir, 2004), an efficient voice treatment for PD patients, would reduce parkinsonian facial movement deficits (Hypothesis I) and result in more positive perceptions of PD patients' personality and behavior (Hypothesis II).

Fifty six participants -- 16 LSVT patients with PD, 12 articulation treatment (ARTIC) patients with PD, 17 untreated PD patients, and 11 demographically-matched controls without PD -- produced monologues about happy emotional experiences on two occasions: Time 1 and Time 2. LSVT and ARTIC were administered during a one-month period between the two occasions. The monologue production task was adapted from the New York Emotion Battery

(NYEB; Borod, Obler, & Welkowitz, 1992). Healthy adult observers (n=110) rated participants' personality and behavior based on participants' videorecorded facial movement. The Facial Action Coding System (FACS) was used to examine changes in the quantity and variability of facial movement (AU Lability Δ and AU Variability Δ, respectively) and complexity of facial expression (AU Complexity Δ).

The increase of LSVT patients on a canonical variate of AU Lability Δ, AU Variability Δ, and AU Complexity Δ was significantly greater than that of ARTIC patients. Additional analyses showed that this result was due to increases in AU Lability Δ and AU Variability Δ of LSVT patients. The personalities of LSVT patients and non-PD controls were rated significantly more positively by observers viewing video clips recorded at Time 2, relative to those recorded at Time 1. Changes in the examined facial movement parameters of LSVT patients did not mediate changes in observers' ratings of those patients.

These findings suggest that LSVT reduces facial movement deficits in PD and possibly results in a more positive perception of LSVT patients' personalities. Results are discussed in the context of studies showing the psychosocial impact of PD patients' communication problems and preliminary evidence regarding the mechanisms underlying LSVT's effect.

Acknowledgements

My work on this dissertation has been a long journey that would not have been possible without the help of quite a few people. I owe my deepest gratitude to my doctoral adviser, Joan Borod, Ph.D., who has given me tremendous intellectual freedom to pursue a wide range of research ideas. Always treating me with incredible kindness, sensitivity, and patience, Joan has helped me overcome many obstacles, supporting me every step of the way. Most importantly, Joan has been a wonderful role model whose incredible devotion to her work has inspired me year after year.

Other committee members, Dr. Harriet Oster and Dr. Laura Rabin, made sure that I learned as much as I could while working on my dissertation. I am extremely grateful to them for their valuable ideas, criticisms, advice, and editing. I also thank Dr. Yaakov Stern and Dr. Yvette Caro whose insightful comments helped me further improve this manuscript.

I am truly indebted to my colleagues, Dr. Lorraine Ramig -- who allowed me to use her dataset for my dissertation -- and Jennifer Spielman, whose helpful and friendly attitude made it a real pleasure to work with her. I am grateful to David McCabe, a fellow graduate student, who joined forces with me on this project. David's involvement and valuable ideas made this dissertation study less daunting and more fun. I have also been very fortunate to work with three great research assistants -- Remington Stafford, Anila Persaud, and Arthur Alimov – who collected personality and behavior ratings: their hard work and dedication have been indispensable. I am also grateful to Alla Landa for helping and inspiring me during the last stages of the process. In conclusion, I most want to thank my parents and my brother. Their love, support, and faith in me during this journey have been important to me in more ways that I can express or even imagine.

Table of Contents

Tables

Figures

Introduction

To familiarize the reader with current knowledge about facial movement deficits in PD, a review of relevant findings is presented. Part I describes the negative consequences of disordered facial communication in PD, showing that these deficits have a negative impact on the perceptions of PD patients by others and thus need to be treated. To examine how facial communication deficits manifest in PD, Part II reviews research on the nature of parkinsonian deficits in spontaneous and posed facial communication. Part III discusses the evidence showing that LSVT, a treatment designed to alleviate voice deficits in PD, also has the potential to reduce PD impairments in facial communication.

I. Social Consequences of Disordered Facial Communication in PD

Patients with Parkinson's disease (PD) demonstrate a variety of impairments expressing themselves through the face and voice. The negative effect that these deficits have on the way PD patients are perceived has long been documented. The fact that even an experienced clinician may initially perceive a PD patient to be aloof or hostile was noted by Monrad-Krohn (1957) more than half a century ago.

The first studies to demonstrate that these clinical impressions partly arise via non-verbal communication cues were carried out by Pentland and colleagues. Pentland, Pitcairn, Gray, and Riddle (1987) showed that after viewing silent videotapes of conversations by PD and cardiac ischemia patients, both students and experienced healthcare professionals, who were blind to patient status, perceived PD patients as more anxious, tense, hostile, suspicious, and unhappy than cardiac patients. PD patients were also seen as more introverted, sensitive, passive, and less intelligent. The substantial difference in the way the two patient groups were perceived stood in stark contrast to the similarity of their scores on self-administered measures of mood,

personality, and intelligence. Another study showed that despite having been trained to understand non-verbal communication, speech therapy students were subject to the same bias (Pentland, Gray, Riddle & Pitcairn, 1988).

These findings raised the question of which non-verbal cues distinguish PD patients from healthy adults and contribute to the negative impression of this patient population. The initial study (e.g., Pitcairn, Clemie, Gray, & Pentland, 1990) to explore this issue compared videotapes of PD patients and a clinical control group with regard to the frequency of various non-verbal cues. The PD patients demonstrated decreased frequency of various communication cues, including several kinds of gestures (e.g., "demonstratives" -- gestures indicating size or distance), postural shifts, and eye movements (i.e., looking away). However, because the perceptions of PD patients by observers were not assessed in the same study, it was unclear whether these group differences in communication behaviors contributed to the negative bias, documented in the earlier study (Pentland et al., 1987), in how PD patients are perceived by other people. Thus, these initial findings did not clarify what, if any, role the observed differences in communication behaviors played in how PD patients were perceived by others. Also, because the videotapes of PD patients in these early studies were silent, the observers did not have access to the information that they would obtain in a conversation with a PD patient. Consequently, this line of research left open the possibility that verbal communication would diminish the negative impression PD patients made on clinicians.

Subsequent investigations (Lyons, Tickle-Degnen, Henry, & Cohn, 2004a; Lyons, Tickle-Degnen, Henry, & Cohn, 2004b; Lyons, Tickle-Degnen, & DeGroat, 2005; Tickle-Degnen & Lyons, 2004) have addressed these methodological limitations. These studies have used a modified version of Brunswik's (1955) lens model of interpersonal perception. The

modified model (see Figure 1) hypothesizes relationships among three constructs: patient personality, clinicians' impression of patient personality, and behavioral cues. According to this model, the correlation between patients' scores on a personality attribute as assessed by a self-administered personality questionnaire and the presence of a behavioral cue (i.e., the relative frequency, intensity, or duration of a cue in the patients' behavior) is a measure of the validity of this cue as an indicator of that attribute (this correlation is denoted "r_v"). The correlation between the presence of a behavioral cue and observers' (i.e., clinicians') ratings of a personality attribute in the patients indicates the degree to which the observers use this cue to make inferences about this attribute (denoted as "r_i"). Finally, the correlation between the observers' ratings of a personality attribute and the patients' scores on this attribute is an estimate of the accuracy with which the observers perceive this trait in patients (denoted as "r_a"). In contrast to earlier studies, observers in these investigations not only saw but also heard the patients describe their favorite activities or challenges of daily living.

Despite their drawbacks (e.g., reliance on self-report of personality attributes), the studies by Tickle-Degnen and colleagues have enriched our knowledge regarding the bases of personality judgments about PD patients. For example, Lyons et al. (2004a) showed that among PD patients, there are accurate behavioral indicators for each of the broad personality dimensions that comprise the validated five-factor model of personality (e.g., Saulsman & Page, 2004): openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. In addition, the cues used for personality judgments by both novice and expert clinicians were facial (e.g., brow furrowing) and vocal (e.g., vocal legibility), and their validity varied depending on the personality trait that was being judged. Though expert clinicians -- whose judgments were invariably more accurate than those of novice clinicians (Tickle-Degnen & Lyons, 2004) --

estimated patients' openness to experience, conscientiousness, and agreeableness with moderate accuracy, all clinicians' judgments of neuroticism and extraversion had little or no relationship to patients' performance on a validated personality questionnaire (Lyons et al., 2004b, 2005; Tickle-Degnen & Lyons, 2004). In contrast to openness and conscientiousness, judgments of neuroticism and extraversion by both novice and expert clinicians were based on the amount of facial masking they observed. Specifically, patients with decreased facial mobility were deemed to be less extroverted and more neurotic. Importantly, it is clear that these findings are consistent with the earlier findings by Pentland and colleagues (Pentland et al., 1988) because most of the personality and behavior characteristics that were perceived to distinguish PD patients from control patients in this earlier study are closely related to neuroticism (e.g., anxious, tense). Moreover, Pentland et al. also found that PD patients were perceived to be more introverted.

Notably, when reported (Lyons, Tickle-Degnen, Henry, & Cohn, 2004b), group means and standard errors indicate that clinicians' perceptions of neuroticism were not only uncorrelated with, but were also significantly higher than, PD participants' self-ratings. Because control groups were not included in these studies, it is not certain whether this difference is a result of the negative bias with which individuals with PD, but not healthy adults, are perceived by others. However, given the results of earlier studies in which control groups were included (e.g., Pentland et al., 1987), this might be the case.

Taken together, these studies suggest that abnormalities in facial communication in PD have a negative impact on the quality of interaction between people with PD and others. This situation may be remedied by educating medical staff and the friends and family of people with PD about these findings. Therapeutic interventions targeting the disturbances in facial communication would also be a promising approach. Before the potential for such interventions

is discussed, studies elucidating the deficits in facial communication in PD (Part II) and the possible neural substrates (Part III) underlying these deficits will be considered.

II. Parkinsonian Deficits in Facial Communication

Two types of facial communication. Early clinical observations formed the basis for one of the most important insights into the nature of PD-related impairment in the control of facial musculature (e.g., Monrad-Krohn, 1924). Specifically, clinicians have long noted that PD patients show a marked decrease in spontaneous expression of emotion on their faces despite being able to produce facial expressions depicting an emotion (e.g., a facial expression depicting anger) when requested.[1] Indeed, the diminished spontaneous expressivity of the face, termed "masked facies," has long been recognized as one of the most reliable PD symptoms (Best & Taylor, 1966). This clinical presentation is the converse of facial expression deficits observed in patients who have sustained lesions to the somatotopic face representation in the primary motor cortex (fPMC) or corticobulbar pathways emanating from that region (e.g., individuals with pseudobulbar palsy). Namely, these patients are often unable to move the muscles on the contralateral side of the face to command, but display normal emotional expressions when spontaneously smiling or crying. Based on this and other evidence, it was believed that voluntary facial movements are dependent on pyramidal pathways originating from the fPMC. By contrast, involuntary facial movements occurring as part of a genuine expression of emotion

---

1 In this dissertation, such facial expressions that are voluntarily produced to depict a mental state that may not be experienced are termed "posed facial expressions;" in contrast, facial expressions that are involuntarily produced as a result of an internal experience are termed "spontaneous facial expressions."

were thought to rely on extrapyramidal pathways arising from premotor areas and subcortical nuclei, including the basal ganglia (Monrad-Krohn, 1924; Rinn, 1984).

Similar to these early observations, subsequent studies investigating facial communication deficits in PD have almost invariably examined these deficits in the context of the facial expressions of emotion.[2] Qualifying early clinical insights, these studies have shown that posed facial expressions, which are dependent on voluntary facial movements, are also affected in PD (e.g., Borod et al., 1990; Bowers et al., 2006a). Additionally, clinicians have documented voluntary facial palsy after damage restricted to extrapyramidal pathways (Jox, Bruning, Hamann, & Danek, 2004). Nevertheless, multiple instances of a voluntary and emotional facial palsy[3] reported in the literature convincingly demonstrate that despite a possible overlap, the pathways for voluntary and involuntary facial communication are largely distinct (e.g., Hopf, Muller-Forell, & Hopf, 1992; Ross & Mathiesen, 1998; Weddell, 1994). In keeping

---

2 Notably, both spontaneous and posed facial communication does not always involve or depict emotion. For example, brow elevation at the moments of emphasis in one's speech (Rinn, 1984) and a facial expression depicting mental effort in the absence of such effort are instances of spontaneous and posed nonemotional facial communication, respectively.

3 An emotional facial palsy is an impairment in spontaneous expression of emotion (i.e., involuntary facial expressions of genuine, felt emotion) without significant deficits in voluntary facial movement; a voluntary facial palsy is an impairment of voluntary facial movement without significant deficits in spontaneous expression of emotion.

with this distinction, research findings regarding spontaneous[4] and posed facial expressions in PD are discussed separately below.

Spontaneous facial expressions in PD. Spontaneous facial expressions in PD have been studied by examining silent videotapes of participants producing monologues about their emotional experiences (Brozgold et al., 1998), participating in semi-structured interviews about their lives (e.g., Pitcairn et al., 1990), viewing affectively loaded material (e.g., Buck & Duffy, 1980), and smelling pleasant and unpleasant odors (e.g., Saku & Ellgring, 1992). The findings to date unequivocally support the early clinical observations of PD-related impairments in spontaneous facial expressions. An early study showed that videotaped nonverbal reactions of PD patients to affectively loaded images were rated by observers as less expressive than the reactions of age-matched controls to the same images (Buck & Duffy, 1980). In addition, the observers were less accurate at guessing the category of an image (e.g., "Familiar People") that evoked a given nonverbal reaction when viewing PD patients than controls. However, because both facial and gestural expressions were videotaped, it was not clear whether the group difference was due to parkinsonian impairments in facial movements, gestures, or both. Given that PD patients were not screened for depression, it is also possible that instead of resulting from parkinsonian impairments in facial movements, the group difference in perceived

---

4 Adequate precautions (e.g., a hidden camera or cover story) were not taken to rule out the influence of "social display rules" (Ekman & Friesen, 1969), social norms dictating how emotions should be expressed in social settings, in the studies of spontaneous facial expressions of emotion in PD. Therefore, it is possible that many of the facial expressions observed in these studies were not spontaneous. Nevertheless, for the sake of expediency, the term "spontaneous facial expressions" will be used to describe these expressions.

expressivity was due to abnormal emotional reactions to slides in depressed PD patients. This possibility is likely given the comorbidity of PD and depression (Cummings & Masterman, 1999), with some estimates of depression in PD being as high as 40% (Frisina, Borod, Foldi, & Tennenbaum, 2008; Raskin, Borod, & Tweedy, 1990).

More convincing are results of subsequent studies that have controlled for differences in depressive symptomatology, mood, or subjective reactions to the presented stimuli between PD and controls (Pitcairn et al., 1990; Simons, Ellgring, & Pasqualini, 2003; Simons, Pasqualini, Reddy, & Wood, 2004; Smith, Smith, & Ellgring, 1996). Two of these studies (Simons et al., 2003; Smith et al., 1996) used the Facial Action Coding System (FACS; Ekman & Friesen, 1978). In this system, any facial activity is coded as a combination of one or several of 44 possible facial actions (e.g., nose wrinkling), termed "action units" (AUs), which are thoroughly described in the FACS manual (Ekman, Friesen, & Hager, 2002b). AUs that can be observed at the same time are termed "AU configurations." For example, if AUs 1, 2, and 9 are observed at the same moment (e.g., on the same frame of a video clip), it is said that a configuration of AUs 1, 2, and 9 (denoted "AU 1+2+9") was produced at that moment. Though some AUs cannot occur together, number of possible AU configurations is extremely high. Further, intensity of most AUs is coded on a 5-point scale.

Studies using FACS have shown that spontaneous expressions of emotion in PD are characterized by decreased facial mobility and decreased intensity. Specifically, these studies have demonstrated that PD patients produce AUs of reduced intensity as they smile (Smith et al., 1996), fewer AUs in a given AU configuration (Simons et al., 2003), fewer changes of facial expression (Smith et al., 1996), and a lower number of smiles with a cheek raise (Smith et al.,

1996) which, in contrast to smiles without a cheek raise, usually occur during spontaneous, but not posed, expressions of emotion (Ekman & Friesen, 1982).

Methods involving raters using subjective criteria[5] to score the quality of facial communication in PD (e.g., its perceived expressivity, number of "false" smiles, perceived amount of facial movement, etc.) have provided confirmatory support for the findings obtained via FACS. Thus, in a recent study, PD patients were perceived to be less expressive than age-matched controls (Simons et al., 2004).[6] Similarly, the results of the FACS analysis showing that individuals with PD, as compared to controls, display fewer smiles with a cheek raise (Simons et al., 2003; Smith et al., 1996) were also corroborated by subjective ratings in a study by Pitcairn and colleagues (Pitcairn et al., 1990). In this study, adult raters perceived PD patients relative to controls to display a lower number of "happy" and a higher number of "false"

---

5 In contrast to the users of FACS and similar systems that code facial movement based on explicit criteria, scores given by naive raters reflect the holistic gestalt-like perception of the human face and are based on less well-defined criteria. For this reason, FACS and similar systems have been termed "objective" measures of studying facial movement, whereas the human ratings are considered to be "subjective" assessments of facial communication.

6 Interestingly, the expressivity levels of PD patients fluctuated depending on the situational context (Simons et al., 2004). Thus, similar to controls, PD patients showed higher expressivity levels when talking to the stranger than when talking to their spouses. Also, their expressivity levels while watching amusing videos with the experimenter were higher than while watching them with a spouse or alone. These findings are encouraging in that they suggest that although the overall level of facial expressivity in PD is dampened, the potential to signal socially relevant factors may be preserved.

smiles. As suggested by studies of smile perception (Frank, Ekman, & Friesen, 1993; Gosselin, Perron, Legault, & Campanella, 2002), the finding by Pitcairn et al. (1990) may correspond, in more objective terms, to a lower number of smiles with a cheek raise and a higher number of smiles without a cheek raise in individuals with PD.

While the reasons for this finding are unknown, a plausible explanation is that after becoming aware of their deficit in spontaneous facial expression in the course of illness, PD patients learn to pose facial expressions whenever an emotional state is experienced. In support of this interpretation, the total number of smiles was similar across the PD and control participant groups in all three studies (Pitcairn et al., 1990; Simons et al., 2003; Smith et al., 1996), suggesting that posed facial expressions may function as a substitute for spontaneous facial expressions in PD. This interpretation implies that individuals with PD are at least partly aware of their deficit in facial mobility and contrasts with a recent finding that PD patients rate their facial expressivity as highly as controls (Simons et al., 2004). However, the Berkeley Expressivity Questionnaire (Gross & John, 1995, 1997) and Interaction Rating Scale (Simons et al., 2004), the expressivity questionnaires administered by Simons et al., did not distinguish between spontaneous and posed expressivity. Thus, it is possible that this finding reflects the PD patients' belief that their compensation strategy is effective (i.e., the belief that by posing facial expressions, they achieve a normal expressivity level) rather than unawareness of their deficit in spontaneous facial expressivity.

Further, it is likely that posed smiles contribute to the negative perceptions of individuals with PD discussed earlier (see Part I) which implies that the compensatory strategy achieves the opposite of its intended effect. The negative impression arising from the increased proportion of posed smiles may also contribute to the perception of normal adults that the nonverbal

communication of PD patients conveys less positive and more negative emotion than the communication of age-matched controls (Brozgold et al., 1998).

As another study (Katsikitis & Pilowsky, 1988) suggests, deficits in spontaneous smiling may be compounded by the hypometria (i.e., decreased movement amplitude) of posed smiles in PD. Specifically, this study revealed a lower degree of mouth opening during posed smiles in individuals with PD relative to controls. Although this result should be considered preliminary as it was not confirmed in a subsequent study (Katsikitis & Pilowsky, 1991), it suggests that posed facial expressions are affected by PD -- an issue that will be considered next.

Posed facial expressions in PD. The impression emerging in the course of a clinical examination that spontaneous, rather than posed, facial expression bears the brunt of neurological damage caused by PD has not been systematically evaluated. Only one study (Smith et al., 1996) compared posed and spontaneous facial expressions on the same dependent variables in the same individuals with PD. This investigation indicated that in contrast to spontaneous smiles, AUs during posed smiles and intensity of these AUs did not differ between control and PD patients and were not affected by disease progression. In contrast to these results, the findings of all other relevant studies, albeit lacking consistency, have shown at least one deficit in posed facial expressions in PD. Taken together, these findings suggest that posed facial expressions are affected in PD, but to a lesser extent than spontaneous facial expressions.

In most of these studies, participants were asked to imitate a presented facial expression or pose an emotion to oral command (e.g., “Look sad”) while they were being videotaped. Using objective measures and subjective ratings, most studies have examined the videotaped expressions in terms of the accuracy or intensity with which a posed emotion was conveyed. For accuracy, most studies have revealed decreased accuracy of raters when judging which emotion

was being posed by PD patients relative to controls (Borod et al., 1990; Madeley, Ellis, & Mindham, 1995; Simons et al., 2004; cf. Borod et al., 1989; Jacobs, Shuren, Bowers, & Heilman, 1995). For intensity, raters in nearly all studies judged that the intensity is reduced in individuals with PD (Borod et al., 1989; Jacobs et al., 1995; Simons et al., 2004; cf. Borod et al., 1990). In addition, applying the method combining FACS with subjective ratings, Simons et al. (2003 & 2004) found that on the task requiring the imitation of a given facial expression, the discrepancy between a participant's face and the target face was significantly larger for PD patients than for controls.

Extending previous findings a step further, Simons et al. (2003 & 2004) also tested PD patients' ability to mask and intensify spontaneous facial expressions. The fact that this ability is needed outside the laboratory to follow social display rules (Ekman & Friesen, 1969) speaks to the ecological validity of these tasks. In these studies, PD patients relative to controls demonstrated reduced overall facial mobility on masking tasks, as shown by subjective ratings (Simons et al., 2004), decreases in average number of AUs in an AU configuration (Simons et al., 2003), and decreases in the range of AUs produced throughout the study session (Simons et al., 2003). These findings show that besides possibly appearing insincere while expressing spontaneous emotions, people with PD have trouble masking them according to social display rules. They also suggest that similar to limb movement, facial movement in PD is subject to akinesia (i.e., decreased ability to initiate movement).

Few studies have directly compared the extent of PD-related impairment in depiction of one emotion type versus another. However, the examination of participant performance in the relevant studies does permit preliminary conclusions regarding the emotions whose facial expression is especially impaired in PD. Specifically, posing disgust, surprise, and happiness, as

compared to other emotions, has nearly always resulted in the largest impairments in PD patients relative to controls (Madeley et al., 1995; Simons et al., 2003 & 2004). Similar to posed facial expressions, spontaneous facial expressions of happiness and disgust, relative to other emotions, were especially impaired (Smith et al., 1996). On the other hand, the posing of sadness, in contrast to other emotions, has consistently resulted in less impairment in PD patients than in controls (Madeley et al., 1995; Simons et al., 2003 & 2004).

The available evidence suggests at least two reasons that the facial expressions of happiness and disgust appear to be especially impaired in PD patients. First, the basal ganglia -- the brain area whose function is severely disrupted by PD -- has been implicated in the processing of these emotions. Specifically, the meta-analysis of 55 neuroimaging studies (Phan, Wager, Taylor, & Liberzon, 2002) showed that this brain region is preferentially activated during the processing of happiness and disgust, as compared to other emotions. Thus, abnormal processing of happiness and disgust in PD may result in the impaired spontaneous expression of these emotions. Second, because the facial expressions of happiness involve significantly greater numbers of facial movements (Smith et al., 1996) and are more accurately identified by raters (Madeley et al., 1995) than other emotional expressions, it is possible that parkinsonian deficits in the expressions of happiness, relative to those of other emotions, are more apparent both in studies using FACS and studies involving human raters. By the same token, the substantial difficulty of both healthy individuals and PD patients when posing sadness (Madeley et al., 1995) may mask PD-related deficits in the expression of this emotion. Clearly, more research is needed to confirm and clarify the differential degree of PD-related impairment in the facial expression of specific emotion types.

Parkinsonian facial expression deficits: the overlap of movement and emotion.

Given that research participants were presented with affectively loaded stimuli or asked to re-experience emotional situations in many studies of facial expression in PD, it is possible that the observed impairments partly resulted from deficits in the perception and/or experience of emotional information that occur in this disease (for review, see Zgaljardic, Borod, Foldi, & Mattis, 2003). As noted earlier, some investigators (e.g., Simons et al., 2004) attempted to rule out this possibility. However, three recent studies have shown that although PD and controls rate the valence of emotional pictures (i.e., the extent to which these stimuli are perceived to be positive or negative) similarly, individuals with PD find them to be less arousing than do controls (Bowers et al., 2006b; Drago, Foster, Skidmore, Trifiletti, & Heilman, 2008; Wieser, Muhlberger, Alpers, Macht, Ellgring, & Pauli, 2006). Because the arousal levels of PD patients were not assessed in the studies of facial expression, it is possible that diminished arousal in this population contributed to reduction of facial mobility. A study by Weddell (1994) also suggests that facial movement deficits cannot fully explain the poverty of emotional expression observed after damage to the basal ganglia. In the Weddell study, spontaneous facial expression of participants with damage to various brain areas was videotaped as they received negative feedback on their test performance. The facial expressivity of participants with damage to the basal ganglia, as compared to other brain regions, remained reduced after any group differences in facial movements during nonemotional speech were statistically removed.

On the other hand, several pieces of evidence suggest that emotional processing deficits cannot fully account for the parkinsonian impairments of facial expression. First, these impairments are evident when overall expressivity (e.g., facial mobility) that includes nonemotional facial expressions (Smith et al., 1996) or when nonemotional expressions in

isolation (Simons et al., 2004) are examined. Second, as discussed earlier, individuals with PD also show deficits in posed facial expressions which are not produced as a direct result of an emotional experience.[7] Third, emotional processing deficits of PD patients appear to be limited. Thus, despite having diminished levels of arousal when processing emotional information (Drago et al., 2008; Wieser et al., 2006), individuals with PD in another study (Halfacre, Borod, Pick, Krch, & Gruber, 2006) produced monologues about their personal experiences with a speech content that was more emotional (i.e., as judged by raters) than the content of monologues produced by healthy controls. Fourth, even when individuals with PD experience emotional processing deficits, such as diminished levels of arousal, it is uncertain whether these deficits translate into reduced facial or verbal output. In one study (Crucian et al., 2001), people with PD, as compared to age-matched controls, were found to use more words when describing experiences of arousing emotions (fear and anger). PD deficits in the speed of cognitive processing or word fluency could not account for this finding as this effect was not obtained for non-arousing emotions or nonemotional experiences. Finally, results of a recent study showed that voluntary facial movements in PD (Bowers et al., 2006a) are characterized by bradykinesia. Namely, as posed facial expressions are produced, they reach peak intensity more slowly in PD

---

7 This dissociation between posed facial expression and emotional experience was observed in patients who had basal ganglia damage caused by tumors or vascular accidents (Weddell, 1994). The posed facial expressions of these patients were characterized by a decreased number of AUs. However, this decrease was completely accounted for by the decrease in number of AUs on another task, the production of nonemotional speech. This finding suggests that facial movement deficits observed on the posed expression task were nonemotional in nature.

patients than in controls. In combination with the other results (Katsikitis & Pilowsky, 1988; Simons et al., 2003) discussed earlier, these findings show that akinesia, bradykinesia, and hypometria -- symptoms affecting the limb movements of individuals with PD -- may also contribute to PD patients' deficits in facial expression.

Taken together, these data suggest that the impairments of facial communication in PD arise from abnormalities in both emotional processing and movement. Because, as argued below, LSVT enhances vocal characteristics and, possibly, facial expression in PD by targeting neural operations that underlie movement, this study focuses on the role of movement, as opposed to emotion, in facial expression. The possibility of improving facial communication in PD patients by addressing emotional processing deficits has important clinical implications and should be tested in future studies.

In order to understand the mechanisms via which LSVT may improve facial communication in PD patients, we will first discuss its putative neural substrates. Because these substrates have not been investigated, hypotheses regarding this issue have to be developed based on the relevant evidence from closely related areas. In the following, such hypotheses will be formed based on the evidence from two literatures: the neural basis of facial movement in the intact brain and the neural basis of PD. The neural substrates of facial movements in the intact brain will be discussed first. Because a comprehensive introduction to the neurology of facial movement is beyond the scope of this dissertation, readers interested in these topics are directed to Rinn (1984) and Morecraft, Stilwell-Morecraft, and Rossing (2004).

III. The Potential of LSVT to Reduce Facial Movement Deficits in PD

We hypothesized that LSVT, as compared to other therapies, may be more likely to have a "carryover" effect on facial movement deficits. This hypothesis is based on empirical evidence

suggesting that both facial and vocal movement deficits in PD appear to result from the same abnormality in a motor planning process, termed "scaling." Moreover, this evidence, reviewed next, also suggests that LSVT achieves the reduction of parkinsionian voice deficits by targeting deficiencies in scaling.

Underscaling in facial and phonatory movements in PD.

Limb and facial movements. Several theories regarding what computational operations are performed by the basal ganglia have been proposed. The leading explanations include sensorimotor integration (Galvan & Wichmann, 2008), selection among competing responses (Gurney, Prescott, & Redgrave, 2001a&b; Mink & Thach, 1991), motor learning (Joel & Weiner, 1994), movement sequencing (Brotchie, Iansek, & Horne, 1991), and movement scaling (Alexander, Crutcher, & DeLong, 1990). Not being mutually exclusive, these accounts emphasize different links in the complex chain of processes that underlie movement.

These accounts help understand various phenomena of abnormal limb movement that are associated with the loss of dopamine in the basal ganglia in PD. However, most of them are not well suited for explaining facial movement deficits in PD. For example, because one's own facial movements usually cannot be seen, sensory feedback, a process that plays a central role in sensorimotor integration theories of basal ganglionic functions, is limited to proprioceptive signals in the production of such movements. Proprioceptive feedback, however, has little influence on movement amplitude (Bard et al., 1999; Rosetti, Desmurget, & Prablanc, 1995) and therefore cannot account for the hypometria of facial movement in PD (Simons et al., 2003). Similarly, it is implausible that selection among competing responses plays a major role during spontaneous facial expressions that are markedly impaired in PD. If this were the case, the sets of muscles involved in a given spontaneous expression (e.g., a smile) and the temporal dynamics

of their movements would vary widely, as is the case with limb movements. However, the opposite is true. Both of these parameters are tightly regulated during spontaneous expressions of emotion and distinguish between spontaneous and posed facial displays (Ekman & Friesen, 1982). Also, facial expressions arise from stereotyped muscle contractions that are perfected in childhood or adolescence and therefore do not involve motor learning in adulthood. Neither can abnormal movement sequencing explain PD-related deficits in facial expressions, because many facial expressions that are impaired in individuals with PD require the contraction of only one muscle (Simons et al., 2003). Disgust that is communicated by the contraction of the caput angulare muscle, which wrinkles the nose, is one such expression.

In contrast to the computational operations emphasized by these theories, deficient scaling appears to play an important role in facial movement deficits in PD. In this dissertation, movement scaling is defined as the neuronal computation of movement speed, amplitude, or force during motor planning, and the term "underscaling" is used to denote a motor planning deficit causing one or more of these parameters to be inadequately small during the resulting motor output. Importantly, because underscaling is a deficit in motor planning as opposed to execution, it results in motor output that is inadequate for achieving one's goals (e.g., pushing a door with adequate strength), rather than in a context-free inability to produce motor output with certain properties (Ho, Bradshaw, Iansek, & Alfredson, 1999).

The contribution of underscaling to hypometria and bradykinesia, two characteristics of parkinsonian facial movements (Bowers et al., 2006a), is supported by the finding that the magnitude of the first burst of electromyographic (EMG) activity during short and simple elbow flexion movements increased with movement amplitude in both PD patients and controls; however, the rate of increase in burst magnitude as a function of amplitude was lower in

participants with severe PD (Pfann, Buchman, Comella, & Corcos, 2001). Additionally, the highest values of burst magnitude were similar in both groups, suggesting that the differences between PD and control groups reflected the impairments of motor planning and not execution. The study also showed that whereas increases of movement amplitude were associated with increases in burst duration in controls, burst duration stayed constant in individuals with PD, possibly contributing to bradykinesia. The findings of another study suggest that when the influence of sensory processes (e.g., vision of the moving limb) on movement execution is eliminated or statistically removed, parkinsonian hypometria becomes more evident, suggesting that motor planning deficits, such as underscaling, are a major cause of this symptom (Desmurget, Grafton, Vindras, Gréa, & Turner, 2003).

Given that PD affects multiple brain areas, it is not clear from these findings (Desmurget et al., 2003; Pfann et al., 2001) which brain areas contribute to movement scaling. This issue has been addressed in an imaging study (Desmurget, Grafton, Vindras, Gréa, & Turner, 2004) which showed activations of the posterolateral putamen, a region of the basal ganglia, in healthy adults who were given a cue signaling the required movement amplitude (versus movement direction) on a reaction-time task. Recent imaging studies have revealed more precisely which physical parameters get scaled in the basal ganglia. These findings have demonstrated that neuronal activity in the internal segment of the globus pallidus (GPi) and subthalamic nucleus (STN) increases with the rate of change of movement force (Spraker, Yu, Corcos, & Vaillancourt, 2007; Vaillancourt, Mayka, Thulborn, & Corcos, 2004) while activity in the caudate, putamen, and the external segment of the globus pallidus (GPe) increases with force duration (Vaillancourt et al., 2004). Involvement of the basal ganglia in scaling is also supported by primate studies showing

the relation between the firing rate of pallidal neurons and movement amplitude/velocity (Georgopoulos, Delong, & Crutcher, 1983).

Because the basal ganglia execute motor control via a functionally segregated neural loop, termed the "motor" circuit (Alexander & Crutcher, 1990; Alexander, DeLong, & Strick, 1986), other areas of the motor circuit, such as the supplementary motor area (SMA) and primary motor cortex (PMC), are likely to be major recipients of underscaled motor programs in PD. The available evidence, including the development of parkinsonian bradykinesia and hypometria (Bhatia & Marsden, 1994) and a decrease in the rate of change of force (Aparicio, Diedrichsen, & Ivry, 2005) following lesions of the basal ganglionic "motor circuit", strongly supports this prediction. This evidence is also consistent with the involvement of the SMA and PMC in force production (Vaillancourt, Yu, Mayka, & Corcos, 2007).

Laryngeal movements. One piece of evidence supporting the role of underscaling in hypokinetic dysarthria,[8] which affects most individuals with PD (Mutch, Strudwick, Roy, & Downie, 1986; Streifler & Hofman, 1984), is the finding that if individuals with PD are instructed to speak loudly and clearly, they manage to do so (Ho, Bradshaw, Iansek, & Alfredson, 1999; Ho, Iansek, & Bradshaw, 1999).[9] Given that laryngeal (also known as phonatory) movements (e.g., the extent of vocal fold adduction) are an important determinant of loudness (Scherer, 1991; Titze & Sundberg, 1992), these data agree with the interpretation that

---

8 Hypokinetic dysarthria is a motor speech disorder that includes, but is not limited to, reduced loudness, reduced pitch inflection, hoarse voice, and imprecise articulation.

9 Notably, based on the outcome of many voice therapies with PD patients, this improvement is transient and does not transfer to natural settings (Allan, 1970; Ramig, Fox, & Sapir, 2004).

reduced loudness and reduced variation of loudness in PD (Duffy, 2005; Ludlow & Bassich, 1984) reflect underscaling in the programming of laryngeal movements, rather than a low-level deficit in movement execution (e.g., muscle rigidity).

This interpretation is also supported by the effectiveness of LSVT, a speech therapy that is based on the view that voice disorder in PD reflects the underscaling of neuronal drive to the laryngeal muscles of the speech mechanism (Trail et al., 2005). In order to rescale this drive, LSVT focuses on increasing vocal loudness in PD patients. The authors of one recent review (Yorkston, Spencer, & Duffy, 2003) concluded that among therapies for respiratory and phonatory dysfunction in hypokinetic dysarthria, "the effects of LSVT clearly have been documented with the widest range of outcome measures" (p. xxxii). In another review (Yorkston, Hakel, Beukelman, & Hager, 2007) focusing on speech production systems (e.g., loudness, speech rate), the authors concluded that that the evidence strongly supports the effectiveness of LSVT in treating reduced loudness in dysarthria.

Emerging neuroimaging data (Liotti et al., 2003; Pinto et al., 2004) also support the hypothesis that the reduction of dysarthria in PD is achieved via the rescaling of motor programs. One study (Liotti et al., 2003) showed that LSVT affects regional cerebral blood flow (rCBF) during phonation and paragraph reading in the basal ganglia and SMA, two areas that, as noted earlier, are implicated in movement scaling (e.g., Vaillancourt et al., 2004). Specifically, the rCBF levels in the right putamen of PD patients were significantly greater during phonation than rest after LSVT. This difference between the two conditions was not obtained before treatment. This effect was accompanied by a significant increase in the patients' vocal loudness following treatment. Conversely, SMA activation in PD patients on the phonation task was observed before, but not after LSVT. The same pattern of SMA activation was revealed on the paragraph

reading task. Interestingly, no SMA activation was observed in controls who performed only the paragraph reading task. These results suggest that SMA activation in PD patients may indicate a compensatory mechanism that became unnecessary when processing in the basal ganglia was enhanced by the treatment (Liotti et al., 2003). Importantly, transient experimenter-cued increases in loudness in untreated patients did not result in any changes in rCBF levels. This result shows that LSVT-related rCBF changes in this study reflected higher-order processes, such as motor planning, rather than motor implementation.

The only other study examining the neuronal correlates of reduced dysarthria in PD (Pinto et al., 2004) also showed increased SMA activity in PD patients as compared to controls during speech production and silent articulation. As in the Liotti et al. study, SMA activity level of the patients during stimulation became similar to that of controls after a successful therapeutic intervention. Taken together, these studies suggest that the successful treatments of hypokinetic dysarthria, such as LSVT, target brain regions involved in the scaling of movement.

Additional considerations. Besides LSVT's rescaling of motor parameters which may contribute to both facial and vocal movements, several other factors increase the likelihood that LSVT-related improvement in vocal characteristics carries over to facial movement deficits in PD. One is the substantial degree of neural coupling between laryngeal and facial movements. Such coupling is thought to occur both via common drive to facial and laryngeal motoneurons and mechanoreceptor input between the face and the larynx (McClean & Tasko, 2002). Evidence of coupling includes high correlations between the timing of vocal fold closure and the timing of lower lip movements (Gracco & Löfqvist, 1994) and between vocal intensity and lip velocities (Dromey & Ramig, 1998). Further, significant correlations of lip velocities with vocal intensity and fundamental frequency (i.e., an acoustic characteristic that underlies pitch and is

determined by laryngeal movements) were obtained in 50-86% of participants with normal speech (McClean & Tasko, 2002). The substantial degree of neural coupling suggests that improvements in temporal dynamics of laryngeal movements and loudness (e.g., Smith, Ramig, Dromey, Perez, & Samandari, 1995; see Yorkston et al., 2003 for an excellent review) following LSVT carries over to lip and, possibly, other kinds of facial movement.

Moreover, the considerable amount of time PD patients spend practicing various components of LSVT raises the likelihood of a substantial carryover effect. In addition to its frequent administration (i.e., four times a week for four weeks) by the speech therapist, LSVT also occurs outside the treatment room. Besides being assigned homework, LSVT patients are instructed to speak louder with family members who frequently reinforce these instructions. This intensive administration format may not only be necessary to achieve the desired effect, but also promote the internalization of rescaled amplitude and help overcome the frequent failure of PD patients to generalize skills practiced with speech therapists to other settings (Ramig et al., 2004). Finally, the simplicity of the LSVT format may be particularly suitable for PD patients, many of whom are elderly and demented (Aarsland et al., 2003). Specifically, LSVT is based on overtraining relatively few, rather than training many, speech characteristics (e.g., adequate loudness). Changes in these characteristics are hypothesized to trigger other positive changes across multiple levels of speech.

Evidence to date. To date, only one known study has examined the possibility that LSVT can alleviate facial movement deficits in PD. Specifically, Spielman, Borod, and Ramig (2003) showed that LSVT, but not a control articulation therapy, resulted in a significant improvement in PD patients' facial mobility, as judged by raters blind to patient status. At present, these results need to be considered preliminary because of several methodological

weaknesses of the study. Besides a small sample size, the content of PD patients' videotaped monologues that were viewed by raters was not controlled. Though the depression levels of participants did not change as a result of treatment, it is possible that therapeutic improvements in the voices of the LSVT patients elevated their mood. This effect may have resulted in the patients' choice of more positive monologue topics leading to increased facial mobility. Nonetheless, Spielman et al.'s findings are encouraging regarding the potential of LSVT to alleviate facial movement deficits in PD.

Study aims.

Aim I. Given the evidence supporting the potential of LSVT in the treatment of impaired facial mobility, we hypothesized that LSVT, relative to a control articulation therapy (ARTIC) and to no treatment, would lead, on average, to significant increases in several measures of facial expression: number of AUs in an AU configuration, number of different facial expressions, and total number of facial expression (Hypothesis 1).

In this study, the number and temporal extent (i.e., start and end times) of facial expressions, which in this study were operationalized as facial events,[10] were coded. Because,

---

10 A facial event, as described in the FACS manual, is either the occurrence of one AU or the occurrence of several AUs that emerge in very close temporal proximity to one another. For example, when, after spotting the face of a friend in a crowd, a person smiles, the elevation of the lip corners (AU 12), the wrinkling of skin near the eyes (AU 6), and the parting of the lips (AU 25) often emerge at nearly the same time. The concept of facial event reflects the fact that external and internal stimuli (a friend's face, stomach pain, etc.) frequently cause several AUs to emerge at roughly the same time. The FACS manual provides guidance on how to identify facial events.

unlike separate AUs, facial events frequently convey a person's reaction to a stimulus, a facial event is the recommended unit of analysis (Ekman & Friesen, 1978) when associations between facial activity and communication are examined. A facial event was the unit of analysis in this study which, as discussed next, investigated the impact of facial movement deficits on patients' communication with observers.

Aim II. Studies discussed in Part I suggest that observers attribute a number of negative characteristics to PD patients, and that these attributions are based on the amount of facial masking observed in a patient (Tickle-Degnen & Lyons, 2004). Based on these findings, we predicted that by increasing facial mobility, LSVT would improve others' negative perceptions of PD patients. These predictions could be formulated as three separate hypotheses.

First, we predicted that the perceived differences between the personalities and behavior of PD participants and control participants obtained in earlier studies (e.g., Pentland et al., 1988) would be replicated in this study (Hypothesis 2a). Because it is possible that a group difference on a given personality or behavior scale would not be obtained, we hypothesized PD patients would be perceived more negatively than controls on aggregate measures of participants' personality and behavior (i.e., termed Overall+, Pers+, and Beh+ scores; see the "A Priori Analyses" section below for the detailed description of how these measures were formed).

Second, we hypothesized that LSVT, as compared to ARTIC or to no treatment, would decrease the negative perceptions and dislike of PD patients observed in earlier studies (henceforth, termed "negative perceptions" for brevity). This decrease would be reflected in the reduced Overall+, Pers+, and Beh+ scores of PD patients (Hypothesis 2b).

Finally, we hypothesized that increases in one or several facial mobility parameters (i.e., average number of AUs in an AU configuration, number of different facial expressions, and total

number of facial expressions) after LSVT would partially or fully mediate the effect of LSVT on the negative perceptions of PD patients (Hypothesis 2c).

The mediation model underlying Hypothesis 2c is shown in Figure 2. Given that there is no available evidence for this model, it was decided to maximize the likelihood of revealing an interesting, if preliminary, support for the model at this exploratory stage, rather than seek only the strongest type of evidence. Thus, as shown in the box on the left of Figure 2, it was decided to operationalize the effect of LSVT as the difference between LSVT and no treatment, rather than LSVT and articulation treatment. Whereas partial mediation would result in the significant coefficients of paths a, b, and c', full mediation would amount to statistical significance of the coefficients of paths a and b only. Notably, it is possible to find support for Hypothesis 2c but not Hypothesis 2b. This would occur if the direct effect of LSVT on personality and behavior attributions counteracted its mediated effect (i.e., path c' and path a x b have coefficients that are opposite in sign) making the overall LSVT effect negligible. Although such suppression effects are inconsistent with Hypothesis 2b, the possibility of their occurrence cannot be ruled out.

Method

Observers.[11]

To assess the effect of LSVT on the perception of PD patients' personality and behavior, 110 observers were recruited from introductory psychology classes. Observers received partial course credit for 1-1.5 hours of their time.

Measures.

The rating forms. After viewing each segment (see below), observers used the Likert scales ranging from 1 (Not at all) to 7 (Extremely) on the Personailty and Behavior Rating forms (see Appendix) to evaluate participants' personality and behavior, respectively. The Personality Rating form included four bipolar scales assessing judgments on the following dimensions: extraverted versus introverted, independent versus dependent, secure versus anxious, and liked (the participant's personality) versus disliked (the participant personality). The Behavior Rating form included five bipolar scales assessing judgments on the following dimensions: engaged versus bored, relaxed versus tense, happy versus sad, openhearted versus guarded, and liked (the participant's behavior) versus disliked (the participant's behavior). These adjectives were adapted from the study by Pentland et al. (1988). In keeping with the principles of creating valid and reliable rating measures (Woods & Hampson, 2005), the position of the positively valenced pole varied across scales (on the Personality Rating form, it was on the right side for 2 out of 4 scales, and on the Behavior Rating form, it was on the right side for 3 out of 5 scales). In keeping with the same principles, descriptions were provided for characteristics whose meaning

---

11 For the remainder of this dissertation, the word "participants" will be used to refer to the 56 videorecorded individuals of this study (see below), and the word "observers" will be used to refer to individuals rating these participants on personality and behavior scales.

can be ambiguous. These descriptions were constructed from the item content of the International Personality Item Pool (IPIP, 2010). Similar to another study describing the construction of a psychometrically sound personality measure (Woods & Hampson, 2005), up to five components were used in a description in order to make the description clear and comprehensive without making it too long. Also, extreme descriptions were avoided (e.g., "often feels blue" instead of "always feels blue" is used to describe "anxious") where possible.

FACS. In FACS, any observable facial movement is represented as a combination of one or several of 44 possible AUs. Also, intensity of most AUs is coded on a 5-point scale. Criteria for scoring each AU and its intensity are detailed in the FACS training manual (Ekman, Friesen, & Hager, 2002b). The names and descriptions of AUs used in the FACS are shown in Tables 1 and 2. High intercoder reliability in the application of FACS has been demonstrated in several publications (e.g., Ekman & Friesen, 1976; Sayette, Cohn, Wertz, Perrott, & Parrott, 2001).

Materials.

Participants. The dataset in this study was drawn from videorecorded monologues produced by 56 individuals from four groups: 16 PD patients who underwent LSVT (LSVT group), 12 PD patients who underwent control articulation therapy (ARTIC group), 17 untreated PD patients (Untreated group), and 11 age- and gender-matched controls (Control group). These videorecordings were made in the laboratory of Dr. Lorraine Ramig at the University of Colorado at Boulder as part of a larger research project (funded by National Institutes of Health grant R01-DC01150). The procedure for eliciting the monologues for this particular study (see the "Development of Stimulus Materials" section below) was developed collaboratively by the laboratories of Dr. Lorraine Ramig and Dr. Joan Borod at Queens College of the City University of New York. The demographics of the four groups are shown in Table 3. As shown by

analyses of variance (ANOVAs) and chi-square tests, the groups were not significantly different in terms of age, gender, race, or years of education, Hoehn-Yahr stage, or time since diagnosis (all p-values > .05).

Participants in the LSVT, ARTIC, and Untreated groups were administered a battery of neurocognitive, psychological, and quality of life assessments (listed in the Appendix). Participants in the Control group were administered the Mini Mental State Examination (MMSE; Folstein et al., 1975) and the 2nd version of the Beck Depression Inventory (BDI-II; Beck, 1967). Moderate to severe dementia (MMSE scores below 25), severe depression (BDI scores above 24), and any prominent abnormalities in neuropsychological functions were criteria for exclusion. An ANOVA followed by pairwise comparisons showed that the mean BDI score of the Control group was significantly lower than that of any other participant group, $F(3, 52) = 4.89$, $p < .01$, $\eta^2 = .83$. This result is consistent with the comorbidity of PD and subclinical depression (Leentjens, 2004; i.e., because patients with BDI scores indicative of severe depression were excluded from this study, the depressive symptoms of some PD patients were likely insufficient for a clinical diagnosis of depression). No other intergroup differences on the mean BDI scores were significant, all p-values > .05.[12]

---

12 Because BDI scores from 14 to 19 are suggestive of "mild" depression (Beck, Brown, & Steer, 1996), the proportion of participants in each group with BDI scores in this range was compared across groups. There were one, four, three, and zero participants with such scores in the ARTIC, LSVT, Untreated, and Control groups, respectively. The proportion of such participants did not differ significantly across groups, $p > .10$. Similarly, the proportion of participants with BDI scores from 20 to 28, which are suggestive of "moderate" depression

The selection of MMSE scores of 25 or below as a criterion for exclusion was based on the population-based norm for MMSE performance of individuals aged 80-84 who, despite having some college experience, do not have a college degree (Crum et al., 1993). With the exception of one participant, such individuals represented the oldest and least educated subset of participants in the present study. Using the individuals with these characteristics as a reference group for selecting the MMSE cutoff score ensured that the cutoff was not overly stringent for any participants in the current study. MMSE scores of 25 or below are more than one standard deviation below the mean MMSE score of this reference group. One standard deviation below the age- and education-specific mean MMSE score is frequently used as the cutoff score in order to screen out participants with cognitive impairment (e.g., Larrieu et al., 2002; Palmer, Wang, Bäckman, Winblad, & Fratiglioni, 2002).

Treatment. Treatment was administered in the laboratory of Dr. Lorraine Ramig at the University of Colorado at Boulder as part of a larger research project (funded by National Institutes of Health grant R01-DC01150). The two types of treatment in the present study, ARTIC and LSVT, were administered in 16 sessions, with a duration of four 50-minute sessions a week for four weeks. Subjects received no additional speech treatment after the 16 sessions. Both types of treatment stimulated high effort in repeated exercises during the first half of each session and the carryover of high effort to speech tasks, such as reading and speaking, during the second half. The treatment intensiveness, daily homework, daily quantification of treatment variables, and carryover were all emphasized equally in both treatment groups. Two clinicians delivered the treatment to all the patients. Both clinicians administered both forms of treatment

---

(Beck et al., 1996), did not differ significantly across groups, $p > .10$. Each patient group had one participant with such scores, and no control participants had BDI scores in this range.

and were randomly assigned to individual patients. The clinicians worked closely together to ensure consistency and equivalent high effort and motivation across both forms of treatment.

ARTIC focused on increasing orofacial-articulatory movement in order to improve overall articulation. ARTIC patients practiced enunciating by performing tasks that included maximum enunciated consonant-vowel combinations (e.g., “pa,” “ta,” and ”ka”), maximum anterior to posterior tongue movement, maximum functional speech enunciation, and hierarchical speech enunciation. Patients were encouraged to enunciate during these tasks and carry over the same increased articulation patterns during initial enunciation tasks into reading and speaking.

The main aim of LSVT was to maximize phonatory efficiency. To this end, LSVT focused on increasing vocal effort by increasing vocal fold adduction. Exercises, such as pushing the hands together or pushing down or lifting up on the arms of a chair or table while phonating, were used to stimulate increased vocal fold adduction. Patients were encouraged to read and speak using the same loud voice that they generated during sustained phonation. Tasks within the LSVT/LOUD protocol included maximum duration of sustained vowel phonation, maximum fundamental frequency range, maximum functional speech loudness, and hierarchical speech loudness exercises. Throughout each session, patients were encouraged to "think loud." Feedback for adequate loudness was given using a voice light, which lights up at target intensity level, and a tape recorder, which was used for self-monitoring of intensity.

Development of stimulus materials. In order to elicit the monologues, the spontaneous emotional expression task from the New York Emotion Battery (NYEB; Borod, Welkowitz, & Obler, 1992) was used as described by Borod and colleagues (Borod, Tabert, Santschi, & Strauss, 2000; Kazandjian, Borod, & Brickman, 2007; Montreys & Borod, 1998). Specifically,

participants described three recent events from their lives: an event that made them angry, an event that made them happy, and an event that made them sad. Thus, three emotional monologues (henceforth, termed "angry monologues," "happy monologues," and "sad monologues") were produced by each participant. In addition, each participant produced an emotionally neutral monologue recounting a recent trip to a supermarket (henceforth, termed "supermarket monologue"). Before beginning the task, a participant was asked to speak for at least 2 mins. If a participant produced a monologue that was less than 2 mins, he or she was prompted by an experimenter to provide additional details about his/her experience. Each participant in the LSVT and ARTIC groups produced four monologues (happy, sad, angry, and supermarket) before and after therapy, with an interval of five weeks between the two monologue productions. Each participant in the Untreated and Control groups produced the four monologues on two occasions separated by an interval of approximately five weeks. For ease of presentation, the first and second occasions on which the monologues were produced by a given participant will be termed "Time 1" and "Time 2," respectively.

In this study, only segments of the happy monologues (henceforth, "happy segments") were analyzed (i.e., coded with FACS and evaluated by observers). Happy segments were chosen for analyses in this study for the following reason. Based on preliminary work with this dataset, the FACS coder in this study, A.D., estimated that he would not be able to code more than two hours of video data in a reasonable length of time (3-4 months). This estimate was based on the correct prediction that, on average, A.D. would code for 20 hours per week. Thus, due to time constraints, only a subset of the videorecordings of participants' monologues could be coded in this study. Rather than code short segments of every monologue, it was decided that a more advantageous approach would be to code longer segments of the happy monologues only.

The rationale for favoring this approach was that a rigorous test of Hypothesis 2c (i.e., hypothesis that changes in perception of LSVT participants would be mediated by changes in facial movement parameters [i.e., parameters which were measured with FACS]) required that both the FACS coder and observers evaluate (via FACS and Rating forms, respectively) the same dataset. Thus, the approach of coding short segments from many monologues would have required that observers evaluate a participant's personality and behavior based on short (approximately 15 s if segments from all monologues were to be coded) segments. Given that observers were asked to make personality and behavior judgments based on facial movements alone (see below), the information contained in short segments may not have been sufficient for making such judgments.

Because, due to time constraints described earlier, coding entire happy monologues was not feasible, it was decided to extract 60 s segments from the end of the monologues. The end, rather than the beginning or middle, of the monologues was chosen based on the following rationale.

First, there is evidence that the emotional intensity of facial expressions of individuals producing monologues from the New York Emotion Battery (Borod et al., 1992) is greater in the last two-thirds than in the initial third of the monologue (Kazandjian, Borod, & Brickman, 2007). Thus, selecting the final rather than the initial segments of the monologues was less likely to result in floor effects when group or pre- versus post-therapy comparisons were made. Second, initial work with the dataset showed that there were more instances of a participant making long pauses near the middle of the monologue than in the last minute of it. Many of these pauses seemed to have occurred when, despite the requirement to speak for at least 2 mins, a participant finished speaking prematurely and was then encouraged to continue the monologue. In these

instances, the resulting monologue often lasted 3 min or longer and contained a long pause that occurred shortly before the beginning of the last minute of the monologue. Thus, extracting the last minute of all monologues yielded a dataset that was relatively uncontaminated by these artifacts.

One hundred twelve segments (56 segments recorded at Time 1 and 56 segments recorded at Time 2) were extracted using the following procedure. First, a research assistant who was not involved in the coding of this dataset listened to every happy monologue and noted the start and end times of every pause in the monologue production that was longer than 5 s and of every instance when an experimenter became visible in a videorecording. Any video data falling between the noted start and end times were defined as an artifact.[13] Sixteen out of 112 recordings (14%) had an artifact. Second, using a videoediting software, all artifacts were removed. Third, the last 60 s were extracted from the resulting artifact-free recordings. These 112 60-second segments were the dataset of this study.

These 112 segments were split into four sets of 28 segments, so that each observer viewed only one set of 28. For expediency, these segment sets will be termed sets "A," "B,"

---

13 To define an artifact, the duration of 5 s was chosen based on the examination of the audio recordings of similar monologues from a different dataset. As was made evident by participants' comments immediately after many longer pauses (i.e., longer than 5 s), these pauses often occurred when participants found it difficult to continue producing the monologue (e.g., could not recall additional details regarding their experience) and thus were probably no longer engaged in recalling the experience. In contrast, shorter pauses (i.e., shorter than 5 s) often occurred before a participant finished relating the gist of their experience. Many of these pauses seemed to have been produced for emphasis or dramatic effect.

"C," and "D." As explained next, this split was carried out for two reasons: to prevent the same observer from judging the personality of the same participant more than once and to avoid observer fatigue. Given that "personality" is generally understood to endure across times and situations (Sullivan, 1953), observers rating the personality of the same person more than once may be biased to rate it similarly on both occasions. To prevent this bias, every observer in this study judged the personality and behavior of a given participant only once. This precaution required at least two sets of observers, with one set of observers viewing 56 (50%) monologue segments. The other set of observers viewed the other 56 monologue segments. A pilot study showed that an observer would have taken, on average, approximately 2.5 hours to view 56 monologue segments and rate the personalities and behavior of 56 participants. Thus, to avoid observer fatigue that would have resulted from rating for 2.5 hours, both sets of 56 monologue segments were further split into two sets of 28 segments, so that each observer would view only one set of 28 segments for approximately 1 hour and 15 min. Thus, 112 monologue segments were split into four segment sets which were shown to four groups of observers, with each group viewing only one set of 28 segments.

The four segment sets were formed via the procedure described next. This procedure ensured that the four sets were equivalent in terms of two independent variables in the statistical comparisons of this study: participant group (PD patients who underwent LSVT, PD patients who underwent ARTIC, untreated PD patients, and age-matched controls) and Time (Time 1 and Time 2).[14] First, a randomly selected half of LSVT participants was assigned to one group

14 If a group of observers had been predominantly exposed to Untreated participants, they might have habituated to the absence of facial movement in the Untreated participants. Such habituation might have caused observers' personality and behavior judgments of Untreated

(Group 1), and the other half was assigned to the other group (Group 2). Second, random assignment to Groups 1 and 2 was repeated for each of the three remaining subject groups (ARTIC, Untreated, and Control). Third, half of 28 participants in Group 1 were randomly selected. Those segments of their monologues that were recorded at Time 1 were assigned to set A, and those segments that were recorded at Time 2 were assigned to set B. Fourth, the segments from the other half of Group 1 participants were also assigned to sets A and B: those segments of their monologues that were recorded at Time 1 were assigned to set B, and those segments that were recorded at Time 2 were assigned to set A. Finally, via the same procedures, the segments of Group 2 participants were assigned to two different sets, C and D.

Procedure.

Rating sessions. The rating sessions were conducted at Queens College. Two to six college students from an introductory psychology class took part in each rating session. Because processing of emotional faces is strongly affected by handedness (Bourne, 2008), only right-handed students were recruited as observers in order to make our observer sample more homogeneous. Also, because individuals speaking English as a second language might not have comprehended some of the text on the Rating Forms (see below), only individuals who spoke fluent English by age 7 were recruited as observers. Students who did not meet these requirements were automatically excluded from participation in this study via an online system that is used by Queens College students to register for participation in research. In addition, we verified during each testing session that observers wrote with their right hand. Before viewing

---

participants to become more positive over time. Thus, the fact that participants from each subject group and segments recorded at both Time 1 and Time 2 were adequately represented in each segment set may have prevented such habituation effects.

the segments, observers were administered a Demographics and Medical History questionnaire (see Appendix) to document their gender, age, ethnicity, number of years of education, and a history of head injury, neurological disorder, learning disability, psychiatric disorder, and alcohol or substance abuse. Observers who reported a history of these abnormalities were excluded.

Four sets of segments were shown to four groups of observers, with each group viewing only one set of segments. Thus, each observer viewed only one monologue (recorded either at Time 1 or Time 2 but not both) from a given participant. The assignment of observers to one of the four groups aimed to keep the average age, average number of years of education, and distributions of gender and ethnicity roughly similar across the four groups of observers. This ensured that any interactions between participant and observer demographics did not confound the outcome of our statistical analyses. Because it was not possible to assign observers while keeping all their demographics similar across the four groups, group similarity on one demographic variable was prioritized over group similarity on another. Specifically, observer demographics were considered during the assignment process in the following order of importance: gender, ethnicity, age, and education. This order is based on two considerations. First, although temporary differences between the four observer groups in terms of age or education may have emerged during the recruitment process, it is unlikely that these differences would have persisted because all observers were college students and thus relatively homogeneous in their age and, especially, education. Second, the available evidence suggests that personality and behavior judgments in this study might have been affected by the interaction between observer gender and participant gender (Murphy, Hall, & Colvin, 2003) and the interaction between participant ethnicity and observer ethnicity (e.g., Lee & Ottati, 1993). However, given that 89% of participants were Caucasian (see Table 3), it is less likely that the

ethnicity of observers would have interacted with the ethnicity of participants than it is that the gender of observers would have interacted with the gender of participants. For this reason, similarity of observer groups on observer gender was given the highest priority during the assignment process. The demographics of each observer group are presented in Table 4. As shown by ANOVAs and chi-square tests, the groups were not significantly different in terms of age, gender, race, or years of education, all p-values > .05.

The segments in each set (A, B, C, and D) were shown to observers twice in 2 different pseudorandom sequences. The sequences were randomly generated, with the only restriction being that participants from the same group were never shown more than 3 times in a row, and that any 4 adjacent segments in one sequence were never adjacent in the other sequence. During one viewing, observers evaluated participant personality, and during another viewing, they evaluated participant behavior. To minimize sequence and order effects, the order in which observers evaluated participant characteristics (i.e., first, personality and then, behavior or vice versa) and which of the two sequences was shown first was counterbalanced across observers. Although, observers' ratings during the second viewing may have been subject to an exposure effect (i.e., the tendency to perceive a stranger in a more positive light upon repeated exposure), the counterbalancing of the order in which observers evaluated participant characteristics ensured that this effect affected personality and behavior to the same extent. Two segments from a different data set were appended to the beginning of each sequence. Unknown to observers, these two segments served as practice trials, and the ratings of people in these segments were not analyzed. Before beginning to rate the personality or behavior of participants in a segment sequence, observers were read the corresponding instructions (see Appendix).

Observers viewed the segments while seated in a quiet room. The segments were shown on a projector screen hanging on a wall. The projector was controlled via a laptop by a research assistant who was invisible to observers while they were evaluating participants. This precaution precluded the ratings from being subject to experimenter effects. To ensure that the monologue content did not influence the ratings, the sound of the videorecordings was turned off. As soon as all observers finished rating a given segment, the next segment in the sequence was shown.

Variables based on observers' ratings. For each segment, we derived an average rating that combined the ratings on all 7-point scales on both rating forms as follows. The extreme rating on a positive attribute (e.g., "very extraverted") or positive attitude ("strongly liked the person") was converted to "7," the moderate rating on a positive attribute (e.g., "moderately extraverted") -- to "6," and so on. Thus, the extreme rating on a negative attribute or negative attitude ("strongly disliked the person") was converted to "1." For each segment, the resulting ratings were averaged across all scales and across all observers. (This composite average, termed the "Overall+ score,"[15] indicated the extent to which a participant in a given segment was

15 As can be seen from examining the personality and behavior characteristics composing the Overall+ score, most of these characteristics are related to the factor of Neuroticism in the five-factor model of personality. However, the score also included an introverted versus extraverted scale that corresponds to the factor of Extraversion in the same model. Thus, the Overall+ score does not measure observers' impressions on a single personality factor. Instead, it is a combination of dimensions on which PD patients were shown to be perceived negatively in earlier studies (e.g., Pentland et al., 1988). Because this study aimed to examine whether LSVT improves these perceptions (Hypotheses 2a and 2b), rather than

liked and perceived positively (i.e., the higher the Overall+ score, the more positive the perception). Additionally, the same procedure was carried out separately with the three personality scales on the Personality Rating form (extraverted versus introverted, independent versus dependent, and secure versus anxious) and four behavior scales on the Behavior Rating form (engaged versus bored, relaxed versus tense, happy versus sad, and openhearted versus guarded). These procedures yielded composite averages, Pers+ and Beh+, indicating, respectively, the extent to which the participant's personality and behavior were positively perceived.

FACS coding.

Scoring process. A FACS-certified coder, A.D., who was blind to the time of the recording (Time 1 or Time 2) and participant group affiliation, used FACS to score the entire dataset (i.e., the last artifact-free 60 s of each monologue segment). While scoring the dataset, A.D. did not know which segments were selected to be scored by D.M. Facial activity in the analyzed dataset was coded according to the rules in the FACS manual (Ekman, Friesen, & Hager, 2002b) with the sound turned off, so that monologue content did not bias the coder. Because the coding of blinks is a very time-intensive activity, and blinks were not the focus of the present study, they were not coded. Initial work with the dataset showed that some participants produced highly repetitive movements that, by virtue of their frequency and unchanging highly stereotypical dynamics, appeared to be unrelated to the information being communicated (e.g., a slight tightening of the lip corners [AU 14] every time a participant closes her mouth). For this reason, movements were judged to have these characteristics by a coder

---

investigate its effect on any personality factor in particular, the approach of combining the attributes of extraversion and neuroticism in one variable does not present a problem.

were not scored. Before scoring the dataset for this study, coders practiced scoring video data that were not analyzed in this study until adequate intercoder reliability (≥.7) based on the standard FACS formula was achieved.[16]

The standard FACS formula is specified in the Investigator's Guide to the Facial Action Coding System (Ekman, Friesen, & Hager, 2002a) and used in many studies utilizing FACS (Ekman & Rosenberg, 2005): intercoder reliability = (the number of AUs on which both coders agree) x 2 / (the total number of AUs scored by both coders). For example, if a given event was scored "AU 6+7+12" by coder A.D. and "AU 6+12" by coder D.M., the number of AUs on which both coders agreed (AUs 6 and 12) and the total number of AUs scored by both coders (AUs 6, 12, 6, 7, and 12) for that event are two and five, respectively. Thus, for these event scores, agreement index = 2 x 2 / 5 = 0.8. If an event was coded by one but not the other coder, agreement for that event was 0.

Variables based on FACS coding. By using FACS, the values of three variables -- AU Complexity (average number of AUs in an AU configuration), AU Lability (total number of facial events), and AU Variability (number of different facial events[17]) -- were determined for

---

16 The data used for practice scoring were the video segments of participants' sad, angry, and neutral monologues. During this practice phase, coders met after scoring a set of 3-4 segments to resolve disagreements in their FACS scores. If the average intercoder reliability across the segments in a set was below .7, the next set of 3-4 segments would be scored and so on. After scoring 10 video segments of 1-2 min duration each, the average intercoder reliability across the last segment set of .74 was achieved.

17 A hypothetical example may clarify the distinction between AU Lability and AU Variability. Whereas all instances of the same event were counted when calculating AU

each video segment. The choice of the three variables, AU Lability, AU Variability, and AU Complexity was based on two considerations. First, to avoid excessive reduction of statistical power, only three variables were selected. Second, to avoid redundancy, variables that measure facial activity on different scales were chosen. On a macroscopic scale, facial activity consists of facial events that unfold over time and may be interspersed with time periods when no observable facial activity is produced (AU 0). By measuring the total number of facial events and the number of different facial events, AU Lability and AU Variability both assessed the richness of facial activity on this gross scale. Despite this conceptual similarity, each of these two variables measured unique aspects of facial movement and thus was deemed to add valuable information that was not assessed by the other variable. On a more microscopic scale, AU Complexity represents the average number of AUs in an AU configuration. The values of each segment on each of three variables -- AU Lability, AU Variability, and AU Complexity -- along with the Overall+, Pers+, and Beh+ scores, were submitted to a statistical analysis, as described in the Data Analyses subsection below.

Lability, each event was counted only once when computing AU Variability. Thus, if, in a video segment, event "AU 1+2+5" changes to event "AU 1+2+4," then changes to event "AU 1+2+5," and then changes to event "AU 1+2+4," the total number of events in that segment is four. Thus, four is the value of variable "AU Lability" for that segment. The number of different facial events in the segment is two ("AU 1+2+4" and "AU 1+2+5"). Thus, the value of variable "AU Variability" for that segment is "2." It should also be noted that AU 0 (the absence of observable facial activity) was not considered to be a facial event or AU configuration in this study. The definition of AU Lability in this study differs from the definition of variable "Lability" in some studies (e.g., Camras et al., 1998).

Calculation of intercoder agreement. To ascertain adequate intercoder agreement, another FACS-certified coder, D.M., independently coded 24 of 112 coded segments (21%). Studies using FACS (e.g., Simons et al., 2003) commonly determine intercoder agreement based on similar percentages of a dataset. The set of 24 segments used in verifying intercoder agreement were created by a research assistant by randomly selecting three segments of LSVT participants at Time 1 and at Time 2 and then repeating the same procedure with ARTIC, Untreated, and Control groups. After both coders completed scoring the dataset, intercoder agreement was calculated based on coders' FACS scores of the same facial events (i.e., coders' scores that began within 1 s of one another and ended within 1 s of one another) after performing the following three steps.

First, disagreements in which one of the coders scored one continuous event, whereas the other coder scored two identical events separated by a brief gap, were excluded from the calculations of intercoder agreement. Such discrepancies likely resulted from both coders seeing the same event, with one coder judging that the event briefly decreased in intensity, and the other coder judging that the event ended and was quickly followed by an identical event. Because AU intensity was not examined in the present study, such discrepancies were not counted as disagreements in the calculations of intercoder reliability if the gap between the two identical events was less than 1 s. Instead, it was considered that both coders scored two identical events in these instances. Overall, less than 4% of the reliability coder's (D.M.'s) AU scores (20 AU scores out of the total of the coder's 565 AU scores) contributed to such disagreements. Second, AU scores for the entire dataset were grouped into facial event scores (see footnote 10) according to the guidelines in the FACS manual. Third, the scores of facial events that can be very difficult to distinguish during speech were combined into one score before calculating

intercoder reliability. Specifically, the following events were not distinguished when calculating reliability: AU 17 (chin raiser) versus AU 17+24 (chin raiser + lip presser), AU 14 (dimpler) versus AU 14+24 (dimpler + lip presser), and AU 18 (lip puckerer) versus AU 18+23 (lip puckerer + lip tightener). Because the present study hypotheses did not concern the occurrences of specific facial events, not taking into account coders' disagreements regarding the distinctions between these similar events when calculating intercoder reliability appeared justified. Overall, less than 4% of disagreements (five out of the total of 144 disagreements) were due to the similarity of these events. After performing these three steps, four types of intercoder agreement were calculated: the percentage of agreement on the occurrence of an event (i.e., the percentage of events scored by the primary coder that were scored by the reliability coder as well), intercoder agreement in judging the onset and offset times of facial events, overall intercoder reliability, and various indices of intercoder reliability specific to the dependent variables of this study -- AU Lability, AU Variability, and AU Complexity. These agreement measures were calculated as follows.

The percentage of agreement on the occurrence of events was calculated as follows. First, the percentage of events scored by the primary coder that were scored by the reliability coder was calculated for each video segment (e.g., if the reliability coder scored 18 out of 20 events scored by the primary coder in a given segment, this percentage was 90% [18/20 = 0.9] for that segment). Second, the percentages of agreement for each segment were averaged across segments. The resulting average (i.e., the overall agreement on the occurrence of events) shows coders' agreement in scoring facial events, regardless of whether coders' event scores are composed of the same AUs.

The difference between the onset times of the coders' scores of the same event, was determined for each event scored by both coders. Next, the differences between the onset times were averaged across all pairs of events scores. The mean difference between the offset times was calculated using the same procedures. These measures indicate whether facial events that were scored by both coders were judged to begin (and end) at the same time.

The overall intercoder reliability was calculated using the standard FACS formula (see the Scoring Process subsection in the Method section). After calculating intercoder reliability for each segment, these reliability values were averaged to determine overall intercoder reliability. This measure is based on the agreement index of each event score and is a widely reported intercoder reliability measure in studies using FACS (Ekman & Rosenberg, 2005). It reflects both coders' agreement whether a facial event occurred and coders' agreement regarding the AUs composing this event.

Finally, to calculate the agreement indices specific to the dependent variables of this study, it was first determined, separately for each variable, whether the distribution of a given coder's scores (i.e., each score indicating the value of a given segment on a dependent measure, such as AU Lability) met the assumption of normality underlying the use of parametric tests. Because for each variable, the distribution of one or both coders' scores violated the normality assumption, the distributions on each dependent measure were rank-transformed. Next, the intraclass correlation coefficients between the rank-transformed scores of one coder and those of the other coder were determined.

Data Analyses

Preliminary analyses.

Testing assumptions underlying parametric tests. Before comparisons of group means in this study (see the Main Analyses and Additional Analyses subsections) were conducted, it was determined whether the contrasted score distributions met the normality and homogeneity of variances assumptions. To test the homogeneity of variances assumption, a modified O'Brien's procedure (Ramsey, 1994) was performed on the compared score distributions. This procedure has been shown to result in a better control of Type I and II error rates than other procedures for comparing the variances of two independent populations (Ramsey & Ramsey, 2007; Ramsey, 1994). If all distributions met both assumptions, an ANOVA was carried out to compare the group means. If either of the assumptions was not met, a nonparametric Welch-James procedure (Lix & Keselman, 1995) was performed to compare the means.

The Welch-James (WJ) procedure is robust to the violation of the assumption of homogeneity of variances (Lix & Keselman, 1995) and can be used for both repeated- and independent-measures designs (Keselman, Algina, & Kowalchuk, 2001; Lix & Keselman, 1995). Because of its robustness, the WJ procedure, relative to ANOVA, results in a better control of Type I errors and a nearly equivalent control of Type II errors when this assumption is violated. Because, to our knowledge, a satisfactory method for testing this assumption has not been devised for repeated-measures designs (Wilcox, 1990), the WJ procedure is more advantageous than ANOVA for testing Hypothesis 1. In addition, it is better suited for repeated-measures designs with unequal group sizes, such as the ones in this study, than other robust tests that are used to detect significant differences among the means of dependent groups (Keselman, Algina, & Kowalchuk, 2001).

Creating canonical variates. AU Lability, AU Variability, and AU Complexity were used as the dependent variables in the validation of facial movement impairment in the present PD patient sample, and changes in these variables from Time 1 to Time 2 (AU Lability Δ, AU Variability Δ, and AU Complexity Δ) were used as the dependent variables in the tests of Hypothesis 1 (i.e., the hypothesis that LSVT, relative to articulation treatment or no treatment, alleviates facial movement deficits in PD). Pearson correlations between AU Lability and AU Variability (see Table 5) and AU Lability Δ and AU Variability Δ (see Table 6) were high, suggesting that these two variables may measure strongly overlapping aspects of facial movement and provide similar and possibly redundant information. This means that analyzing them separately may also be redundant. To address this issue, the dependent variables were combined into one variable, termed canonical variate, via a discriminant function analysis (DFA). Specifically, as detailed in the Results section, AU Lability, AU Variability, and AU Complexity were combined into a canonical variate CV_FM to validate facial movement deficits in the present PD patient sample. Similarly, AU Lability Δ, AU Variability Δ, and AU Complexity Δ were combined into a canonical variate CV_H1 to test Hypothesis 1.

Main and exploratory analyses.

Because very little is known about the effect of LSVT on facial movement, limiting this study to the examination of few variables or relationships would not be justified. Testing the same hypotheses with a wider range of facial mobility measures and statistical comparisons may provide valuable information for subsequent, more targeted studies. Thus, the tests of Hypotheses 1, 2a, and 2b included both main and exploratory analyses.

Validation of facial movement deficits in the sample of PD patients. Because Hypotheses 1 and 2c of this study are based on previous evidence showing the presence of facial

movement deficits in PD patients, the analyses ascertaining the presence of these deficits were carried out first. Specifically, the facial mobility of 45 PD participants at Time 1 (before treatment) was compared to that of 11 age-matched control participants on each of the four dependent variables: AU Variability, AU Lability, AU Complexity, and CV_FM. To this end, a one-way WJ procedure with Group (PD patients versus controls) as a between-subjects factor was performed, separately for each dependent variable.

Tests of Hypothesis 1.

Main analyses. Hypothesis 1 (i.e., hypothesis stating that LSVT, relative to ARTIC and no treatment, reduces facial mobility deficits) was probed by performing planned intergroup comparisons on four dependent variables: AU Lability Δ, AU Variability Δ, AU Complexity Δ, and CV_H1. Specifically, for each dependent variable, a separate one-way WJ procedure was performed to contrast the mean of three non-LSVT groups (ARTIC, Control, and Untreated) with the mean of the LSVT group (Contrast 1). Similarly, for each dependent variable, a separate one-way WJ procedure was performed to contrast the mean of the LSVT group with the mean of the ARTIC group (Contrast 2). As discussed earlier, these contrasts were designed in order to test as directly as possible whether, as Hypothesis 1 predicts, increases in the facial mobility of the LSVT patients exceed those of the other participants. As the last step, analyses of covariance were performed to determine whether eliminating the effect of extraneous variables (e.g., participant's education level) affected the significant results revealed in Contrasts 1 and 2.

Exploratory analyses. The exploratory analyses of Hypothesis 1 were motivated by three aims. The first aim was to further test Hypothesis 1 by performing additional intergroup comparisons. For instance, Hypothesis 1 predicts that changes (from Time 1 to Time 2) in the

facial mobility of the LSVT patients are greater than those of the untreated PD patients. Because the main analyses of Hypothesis 1 were limited to comparing the groups via Contrasts 1 and 2, this prediction was not tested directly in the main analyses. Testing this prediction on each facial mobility measure provided additional evidence regarding which facial mobility aspects change as a result of the LSVT. In addition, to rule out the possibility that permanent intergroup differences in facial mobility confounded the tests of Hypothesis 1, the relevant intergroup comparisons were performed again, with AU Variability Δ, AU Lability Δ, and AU Complexity Δ measured as a proportion of the participant's overall level of facial mobility rather than in absolute terms (see below).

The second aim was to examine whether the magnitude of facial movement changes in the LSVT group depended on the characteristics of a patient. For instance, it is possible that LSVT had a greater effect on a certain patient segment (e.g., patients at a more advanced Hoehn-Yahr stage).

Finally, the third aim of the exploratory analyses was to replicate the findings of the DFA (discriminant function analysis), a statistical method used to combine facial movement measures into CV_H1, in a more homogenous sample that consisted entirely of PD patients. Such replication would provide additional support that the high coefficients of AU Variability Δ and AU Lability Δ in the CV_H1 formula reflected LSVT-related changes in facial variability and lability rather than the idiosyncrasies of our participant sample.

Tests of Hypothesis 2a.

Main analyses. To test Hypothesis 2a (i.e., the hypothesis stating that observers' negative perceptions of PD patients that were obtained in earlier studies would be replicated in this study), a separate one-way WJ procedure with Group (PD patients versus controls) as a between-

subjects factor was performed on each dependent variable: Overall+, Pers+, and Beh+. To avoid confounding the effects of treatment with those of PD, only Time 1 ratings were entered in the analysis. Next, analyses of covariance were carried out to examine whether eliminating the influence of extraneous variables (e.g., observer gender, Set [i.e., which of the four sets of clips was shown to an observer]) affected the outcomes of intergroup comparisons.

Exploratory analyses. Three sets of exploratory analyses of Hypothesis 2a were conducted. The first set explored which aspects of personality and behavior were perceived to be more negative in PD patients than in others. These analyses aimed to provide a more detailed description of the negative impressions with which this clinical population is perceived. To this end, the raw ratings of PD patients were contrasted with those of controls on each personality and behavior scale. The second and third sets of analyses investigated whether observer and participant characteristics, respectively, played a role in the negative perceptions with which PD patients were perceived. To address this issue, the effect of such characteristics (e.g., gender and time since PD diagnosis) on Time 1 ratings was examined.

Tests of Hypothesis 2b.

Main analyses. To examine whether, as Hypothesis 2b predicts, LSVT, relative to the articulation treatment or no treatment, reduces observers' negative perceptions of PD patients, changes (from Time 1 to Time 2) in observers' ratings of LSVT patients were compared with the ratings of other patients. Namely, changes in the ratings of the LSVT group were contrasted with those of the non-LSVT groups (Contrast 1) via a WJ procedure, separately for each of the three rating types (Pers+, Beh+, and Overall+). Similarly, changes in the ratings of the LSVT group were contrasted with those of the ARTIC group (Contrast 2) via a WJ procedure, separately for each of the three rating types (Pers+, Beh+, and Overall+). Next, analyses of

covariance were conducted to determine whether eliminating the effect of extraneous variables affected the results of these contrasts.

Exploratory analyses. Two sets of exploratory analyses of Hypothesis 2b were conducted. The first set of analyses examined more thoroughly whether variable Time (Time 1 versus Time 2) had an effect on the ratings of each participant group. Because Contrasts 1 and 2 were targeted to identify specific Time x Group interactions, they do not detect the within-group effects of variable Time. Such effects may qualify the lack of significant effects in the main tests of Hypothesis 2b (i.e., hypothesis stating that the LSVT improves the negative impressions of PD patients) and have implications for future studies. To reveal such effects, additional comparisons that were not included in the main analyses were carried out.

The second set of analyses explored whether the magnitude of the LSVT's effect depended on patient characteristics. To investigate this question, the effect of these characteristics on the amount of change (from Time 1 to Time 2) in the ratings of the LSVT group was probed.

Tests of Hypothesis 2c. Based on Hypothesis 2c (i.e., the hypothesis that the effect of LSVT, relative to untreated patients, on the ratings of PD patients is mediated by changes in facial movement parameters), a mediation model, shown in Figure 2, was hypothesized. To test the model, mediation analyses were performed according to the modified six-step procedure recommended by Shrout and Bolger (2002). These researchers showed that by taking into account the violation of normality in the distribution of certain parameters in the mediation model, this six-step procedure is more powerful than the classical approach to mediation analysis that does not account for this violation. The procedure includes the testing of regression equations as well as bootstrap methods and can be conducted in SPSS (Statistical Package for the

Social Sciences) by using the appropriate programming syntax. The procedures performed to test the mediation model of change in participants' Overall+ scores are detailed next. Identical steps were conducted to test the mediation models of changes in participants' Pers+ and Beh+ scores.

Step 1 is to determine whether changes in an independent variable are associated with changes in an outcome variable. In the present model, the independent variable is treatment status (LSVT versus Untreated), and the outcome variable is the Overall+ ratings. To carry out Step 1, the mean Overall+ rating of the LSVT group was contrasted with that of the Untreated group via a WJ procedure. Step 2 is to determine associations between the independent variable and mediators (i.e., facial movement measures). To perform Step 2, the mean of the LSVT group was contrasted with that of the Untreated group via a WJ procedure, separately for each mediator. This step revealed whether paths $\underline{a}_{\text{LSVT-Num}}$, $\underline{a}_{\text{LSVT-Lab}}$, and $\underline{a}_{\text{LSVT-Var}}$, were significant. Paths that were not significant and the mediators that were not influenced via any $\underline{a}$ paths were dropped from the model as suggested by Shrout and Bolger. Step 3 is to determine associations between outcome variables and mediators. To perform this step, changes in the Overall+ scores were regressed on each of the mediators remaining in the model while holding the remaining $\underline{a}$ paths constant. This was done separately for each of the remaining mediators via a nonparametric least trimmed squares regression. Because Step 3 showed that none of the remaining $\underline{b}$ paths were significant, Steps 4-6 were not performed in accordance with Shrout and Bolger's recommendations.[18]

---

18 If Step 3 revealed at least one significant b path, Step 4 would be to test the significance of the mediated effect via the remaining *a* and *b* paths by using bootstrap methods. Step 5 would be to determine whether the direct effect of treatment on the outcome variable is

Results

Analytic Strategy

The results of this investigation are detailed in the next five sections. The first section presents various indices of intercoder reliability. Because Hypotheses 1 and 2c of this study are based on previous evidence showing the presence of facial movement deficits in PD patients, the second section describes analyses aimed to ascertain the presence of these deficits in the present PD patient sample. The third section describes the results of analyses testing Hypothesis 1, which predicts the alleviation of facial movement deficits as a result of LSVT. The fourth section reports the results of analyses examining Hypotheses 2a and 2b. Finally, the last section details the results of analyses probing Hypothesis 2c.

Given the exploratory nature of this study, it was decided to thoroughly examine this dataset rather than to limit the analysis to few statistical comparisons. Thus, besides the planned comparisons discussed in the Method section, a number of exploratory analyses were performed. For this reason, where appropriate, the sections are divided into the "Main Analyses" and "Exploratory Analyses" subsections. As the goal of the exploratory analyses was to provide preliminary findings for replication in future studies, the statistical significance levels of these analyses were not adjusted for multiple comparisons. Such adjustments would have drastically reduced statistical power making the discovery of potentially important findings unlikely.

---

significant (path *c'* in the model) by regressing the outcome variable on the independent variable while holding any significant *b* paths constant. If the effect of treatment status on Overall+ scores were mediated by a change in facial movement (i.e., one or more *a* x *b* paths in Step 4 are significant), and there were no strong evidence of suppression effects, Step 6 would be to assess the strength of mediation effect. See Shrout and Bolger (2002) for further detail.

Because these adjustments were not made, the Type I error rate for the exploratory analyses is considerably higher than its conventional level of 5%. Thus, the results of these analyses will be interpreted most cautiously.

Intercoder Agreement

Overview of analyses. Before calculating the indices of intercoder agreement, procedures detailed in the Calculation of Intercoder Agreement subsection (see the Method section) were performed. For all intercoder agreement calculations, coders' scores that began within 1 s of one another were considered to be the scores of the same facial events. This section reports four measures of intercoder agreement: the percentage of agreement on the occurrence of an event (i.e., the percentage of events scored by the primary coder that were scored by the reliability coder as well), intercoder agreement in judging the onset and offset times of facial events, overall intercoder reliability, and various indices of intercoder reliability specific to the dependent measures of this study -- AU Lability, AU Variability, and AU Complexity.

Agreement indices. The percentage of agreement on the occurrence of an event was 81% ranging from 3 to 100% for a given segment. For 50% of participants (12 out of 24), the percentage of agreement on the occurrence of an event was 95% or greater. The average differences between the onset and offset times were 0.17 s (ranging from 0.00 to 0.96 s) and 0.58 s (ranging from 0.00 to 0.98 s), respectively. The average overall intercoder reliability, based on the standard FACS formula for intercoder agreement (see the Scoring Process subsection in the Method section), was .71 ranging from .26 to 1 for a given segment. The overall intercoder reliabilities in scoring the video segments of PD patients ($n = 18$) and normal controls ($n = 6$) -- .73 and .64, respectively -- were not significantly different, $p > .10$, suggesting that the reliability of coding did not depend on whether a participant in a segment had PD. To examine whether the

coding of facial movements produced by advanced PD patients was more or less reliable than that of facial movements produced by PD patients in the early disease stages, two Pearson correlations were computed. First, the overall intercoder reliability values for each segment of a PD patient's monologue (n = 45) were correlated with the Hoehn-Yahr stages of those patients. Second, the overall intercoder reliability values were correlated with time that has elapsed since the diagnosis of each PD patient. Neither correlation was significant, $p > .10$, suggesting that the severity of PD did not affect the reliability of coding. The intraclass correlation coefficients between one coder's scores and those of the other coder were .80, .75, and .60 for AU Lability, AU Variability, and AU Complexity, respectively.

Facial Movement Deficits in the Present Sample of PD Patients

Overview of analyses. This section describes the results of analyses testing whether facial movement was impaired in the present PD sample. To this end, the facial mobility of 45 PD participants at Time 1 (before treatment) was compared to that of 11 age-matched control participants. Three measures of facial movement, that is, AU Lability (the number of facial events in a monologue clip), AU Variability (the number of unique facial events in a monologue clip), and AU Complexity (the average number of AUs occurring simultaneously during a facial event), were examined. As discussed in the Method section, these variables were chosen based on previous studies showing group differences on these measures (e.g., Camras et al., 1998). The means on these three measures at Time 1 and Time 2 are shown in Figures 3-5 and Table 7 for each participant group. Pearson correlations among these variables within each of the two groups (PD patients and control participants) are shown in Table 5. In addition, based on a rationale explained below, we compared the means of PD patients and controls on a linear combination of these three variables, termed a "canonical variate."

Thus, the analyses of parkinsonian facial movement involved three steps. First, the difference between the means of two groups (PD patients and controls) on each dependent variable was tested for statistical significance. Second, three dependent variables were combined into a canonical variate via a discriminant function analysis (DFA). Finally, the difference between the group means on this variate was tested for statistical significance. As shown below, these analyses confirmed that the facial mobility of PD patients was reduced relative to that of controls.

Statistical comparisons of group means on AU Lability, AU Variability, and AU Complexity. To determine whether parametric procedures could be performed to compare the group means, the normality and homogeneity of variance assumptions underlying parametric tests were tested. The modified O'Brien's procedure (Ramsey, 1994, see the Method section) showed that the variance of AU Variability scores was not homogenous across the two groups (PD patients and controls), $\underline{F}(1, 54) = 5.76$, $\underline{p} = .02$. In addition, there were positive skew and kurtosis in the AU Lability and AU Variability score distributions of PD patients ($\gamma_1 = 2.04$, $\gamma_2 = 5.47$ for AU Lability and $\gamma_1 = 1.15$, $\gamma_2 = 2.61$ for AU Variability) and in the AU Complexity score distribution of controls ($\gamma_1 = 1.54$, $\gamma_2 = 1.36$). Thus, a nonparametric Welch-James (WJ) procedure (Lix & Keselman, 1995) was carried out to contrast the mean of PD patients at Time 1 versus that of controls, separately on each dependent variable.[19] Compared to other

---

19 Statistical analyses were carried out to examine whether the means of the PD groups on each dependent variable at Time 1 were similar. Given the violations of the normality assumption in the score distributions on all dependent variables, a WJ procedure was carried out for this purpose, separately for each dependent variable. For each dependent variable, the effect

nonparametric tests, this procedure was shown to result in low Type I and Type II error rates (Roth, 1988), is robust to the violations of normality and homogeneity of variance assumptions (Lix & Keselman, 1995), and can be used for both repeated- and independent-measures designs (Keselman, Algina, & Kowalchuk, 2001).

A WJ procedure revealed that the AU Lability mean of controls (M = 15.27, SD = 10.98) was significantly greater than that of PD patients (M = 7.27, SD = 7.35), $F(1, 12.34) = 5.36$, $p = .04$, ES = 1.15. In addition, the difference between the AU Variability mean of controls (M = 7.00, SD = 4.77) and that of PD patients (M = 3.80, SD = 2.79) was marginally significant, $F(1, 12.09) = 4.50$, $p = .055$, ES = 0.85.[20] The AU Complexity means of the PD (M = 1.51, SD = 0.79) and Control (M = 1.83, SD = 0.97) groups were not significantly different, $F(1, 13.13) = 1.22$, $p = .29$.

These results suggest that in partial support of our predictions, the facial movement of PD patients is impaired on some but not all facial movement variables. Specifically, the lability and, possibly, variability, but not complexity, of facial expressions in PD is significantly lower than the same parameters of facial expressions in controls. However, it may be argued that because the correlation between AU Variability and AU Lability is high in all three datasets, these two variables measure strongly overlapping aspects of facial movement and provide similar and possibly redundant information. This means that analyzing them separately, as was done in the

---

of Group was not significant (all p-values > .10) showing that three patient groups did not differ significantly in terms of facial lability, variability, and complexity at Time 1.

20 The alpha level of .05, two-tailed, was used for testing statistical significance in the present study. However, in keeping with the exploratory nature of many analyses discussed in this dissertation, trends associated with the p-values in the .05-.10 range will also be reported.

preceding contrasts, may also be redundant. To address this issue, three dependent variables were combined into one variable via a discriminant function analysis (DFA) and the means of two groups (PD patients and controls) were statistically compared on the resulting variable.[21]

Discriminant function analysis. The DFA is a statistical procedure for creating new dependent variable(s) on which participant groups -- or, more generally, the levels of an independent variable -- differ the most. Each new variable, termed a canonical variate (CV), is a linear combination of the original dependent variables. Expressed mathematically, $CV = a_1*DV_1 + a_2* DV_2 + ... + a_n* DV_n$, where DV_1, ..., DV_n are n dependent variables and a_1, ..., a_n are coefficients determined via the DFA. The number of CVs produced by the DFA is equal to the number of levels of the independent variable minus one or the number of original dependent variables, whichever is less. Unless a CV is statistically significant, as indicated by the corresponding p-value, one cannot be confident that it provides the best discrimination among the groups.

To calculate the first CV, the entire variance of the dependent variables is used in the DFA. To compute the second CV, the variance that is not captured by the first CV is utilized. Because in the calculation of a CV, only the variance not captured by the preceding CVs is used, CVs produced by the DFA are orthogonal (i.e., statistically independent), and differences among

---

21 Although AU Complexity was not highly correlated with other dependent measures, it was entered in the DFA because it was deemed that the more dependent variables are entered in this procedure, the more accurate its results will be. To rule out the possibility that entering AU Complexity in the DFA confounded the present findings, only AU Variability and AU Lability were entered in the DFA, and all analyses based on the results of this procedure were repeated. All the results of these analyses remained the same.

averages on one CV have no relation to differences among averages on another. Thus, unlike statistical comparisons on highly correlated dependent variables, comparisons on the CVs generally do not provide redundant information.

To mitigate the difficulties presented by the highly correlated dependent variables, a DFA with a between-subjects factor PD Status (PD patients versus controls) and AU Lability, AU Variability, and AU Complexity as three dependent variables was performed to find the CVs providing the best discrimination between PD patients and controls. Only the first CV, denoted henceforth as "CV_FM," was significant and had the following coefficients:

CV_FM = 0.68*(AU Variability) + 0.50*(AU Lability) - 0.30*(AU Complexity).

Guidelines for interpreting canonical variate equations recommend that coefficients with an absolute value of 0.40 or greater be considered high suggesting that variables with such coefficients contribute meaningfully to the resulting CV (Stevens, 1996). In contrast, when the absolute value of coefficients is 0.30 or less, the contribution of the corresponding variables is generally considered trivial. Thus, similar to group comparisons on the original dependent variables (AU Lability, AU Variability, and AU Complexity; see the preceding subsection), the CV_FM equation suggests that higher facial variability and lability of controls, relative to PD patients (see Figures 3-4 and Table 5), provided the best discrimination between the two groups. In contrast, the mean complexity of facial movement (see Figure 5) did not differentiate well between the patients and controls. To confirm the presence of facial movement deficits in PD patients, the CV_FM means of these two groups were statistically compared, as reported next.

Statistical comparison of CV_FM group means. The CV_FM means of PD patients (M = 5.76, SD = 5.19) and normal controls (M = 11.85, SD = 8.56) are shown in Figure 6.

These means were contrasted via a WJ procedure[22] which revealed that the CV_FM mean of controls (M = 11.85, SD = 8.56) was significantly higher than that of PD patients (M = 5.76, SD = 5.19), F(1, 12.05) = 5.11, p = .04, effect size (ES) = 1.03.[23] This finding suggests that the variability and/or lability of facial movement is lower in PD patients than in controls.

In sum, the significant differences between the group means on the original dependent variables (AU Lability, AU Variability, and AU Complexity) were consistent with the group differences on the CV_FM. Both sets of analyses demonstrated that compared to age-matched controls, PD patients showed a decreased range of facial expression (i.e., decreased AU Variability) and changed facial expressions less frequently (i.e., decreased AU Lability). In contrast, the complexity of facial expression did not significantly differ between PD patients and age-matched controls.

Hypothesis 1: LSVT, Relative to ARTIC or No Treatment, Reduces Facial Mobility Deficits

---

22 Nonparametric WJ procedures were chosen for intergroup comparisons because the score distributions compared in these tests violated the homogeneity of variance and normality assumptions. Though the modified O'Brien's procedure showed that the homogeneity of variance assumption was satisfied, $F(1, 54) = 3.23$, $p > .10$, the CV_FM score distribution was positively skewed and leptokurtic in the PD group ($\gamma 1 = 1.60$, $\gamma 2 = 3.33$) violating the assumption of normality.

23 Statistical analyses were carried out to examine whether the CV_FM means of the PD groups were similar. Given the violations of the normality assumption in the CV_FM score distributions, a WJ procedure was carried out for this purpose. The effect of Group was not significant (p-value > .10) showing that the CV_FM means did not differ significantly across three patient groups.

Overview of analyses. To test Hypothesis 1, changes on facial mobility measures from Time 1 to Time 2 were compared across four participant groups. Changes in AU Lability, AU Variability, and AU Complexity from Time 1 to Time 2 (AU Lability Δ, AU Variability Δ, and AU Complexity Δ) were used as the dependent variables in the analyses, and their mean values are shown in Figures 7-9. As shown in Table 6, AU Lability Δ and AU Variability Δ were highly correlated. Thus, the same analytic approach as in the preceding section (see the "Overview of Analyses" subsection of the "Facial Movement Deficits in the Present Sample of PD Patients" section) was followed. Namely, to eliminate the potential redundancy in the information provided by the highly correlated variables, the participant groups were compared not only on AU Lability Δ, AU Variability Δ, and AU Complexity Δ, but also on a canonical variate (CV) derived from these variables. This CV was the linear combination of these variables that allowed the best discrimination of LSVT patients versus other participants in terms of facial mobility changes from Time 1 to Time 2. As noted earlier, the analyses of Hypothesis 1 were divided into the main and exploratory analyses. To prevent inflating the Type I error rate, the main analyses were limited to performing only planned intergroup contrasts (i.e., Contrasts 1 and 2, described next), whereas the exploratory analyses included other relevant comparisons.

Hypothesis 1: Main analyses. The main tests of Hypothesis 1 were carried out in four steps. In Step 1, Hypothesis 1 was probed by performing planned intergroup comparisons on the three dependent variables: AU Lability Δ, AU Variability Δ, and AU Complexity Δ. Specifically, Contrasts 1 and 2 were performed separately on each dependent variable. As the reader may recall, the mean of the LSVT group was statistically compared with the mean of three non-LSVT groups (ARTIC, Control, and Untreated) in Contrast 1, and the mean of the LSVT group was statistically compared with the mean of the ARTIC group in Contrast 2. As

discussed in the Method section, these contrasts were designed in order to test as directly as possible whether, as Hypothesis 1 predicts, increases in the facial mobility of the LSVT patients exceed those of the other participants. In Step 2, a DFA was used to combine AU Lability Δ, AU Variability Δ, and AU Complexity Δ into a CV (canonical variate). In Step 3, the group means on the CV that was obtained in Step 2 were compared via Contrasts 1 and 2. In Step 4, analyses of covariance were performed to determine whether eliminating the effect of extraneous variables affected the findings obtained in Step 3.

As detailed below, in support of Hypothesis 1, the obtained findings suggested that the LSVT patients, as compared to the ARTIC and other participants, demonstrated a greater increase on the facial variability and lability, but not complexity, of facial movement. Moreover, no extraneous variables accounted for these intergroup differences.

Step 1: Intergroup comparisons of AU Variability Δ, AU Lability Δ, and AU Complexity Δ group means. A WJ procedure[24] was performed to compare the group means via

---

24 Nonparametric WJ procedures were chosen for statistical testing because the score distributions compared in these tests violated the assumptions underlying the use of parametric tests. Although the modified O'Brien procedures showed that for each variable, the homogeneity of variance assumption was met (all p-values > .10), five score distributions violated the assumption of normality. Specifically, the AU Lability Δ distribution was negatively skewed and leptokurtic in the Control ($\gamma 1 = -0.85$, $\gamma 2 = 1.37$) and Untreated groups ($\gamma 1 = -1.97$, $\gamma 2 = 7.91$), *and* the AU Variability Δ distribution was negatively skewed and leptokurtic in the ARTIC group ($\gamma 1 = -1.87$, $\gamma 2 = 4.47$). The AU Complexity Δ distribution was negatively skewed and leptokurtic in the ARTIC group ($\gamma 1 = -1.14$, $\gamma 2 = 1.95$) *and* positively skewed and leptokurtic in the LSVT group ($\gamma 1 = 2.36$, $\gamma 2 = 8.15$).

Contrasts 1 and 2, separately on each dependent variable. The results of these comparisons are reported next. Besides Contrasts 1 and 2, other relevant intergroup comparisons were performed to more comprehensively examine the dataset. However, because these additional comparisons were post hoc, their results are reported in the "Exploratory Analyses" subsection below.

Contrasts 1 and 2 performed via a WJ procedure showed that the AU Complexity Δ mean of the LSVT group (M = 0.40, SD = 1.31) was not significantly different from that of the other groups combined (M = 0.12, SD = 0.71), $F(1, 18.72) = 0.48$, $p = .50$, or that of the ARTIC group (M = 0.11, SD = 0.57), $F(1, 24.63) = 0.57$, $p = .46$. For AU Variability Δ, the mean of the LSVT group (M = 1.81, SD = 3.04) was significantly greater than that of the other groups combined (M = -0.78, SD = 2.62), $F(1, 25.52) = 8.48$, $p < .01$, ES = 0.81. The difference between the AU Variability Δ means in the LSVT versus ARTIC (M = -0.58, SD = 3.15) groups was marginally significant, $F(1, 23.38) = 4.09$, $p = .06$, ES = 0.76. For AU Lability Δ, the mean of the LSVT patients (M = 5.00, SD = 8.23) was significantly greater than that of the other groups combined (M = -1.43, SD = 6.87), $F(1, 27.65) = 4.88$, $p = .04$, ES = 0.72. The AU Lability Δ mean of the LSVT group was not significantly different from that of the ARTIC group (M = -1.25, SD = 6.68), $F(1, 24.91) = 2.13$, $p = .16$.

Taken together, these results provide moderate support for Hypothesis 1. Namely, changes (from Time 1 to Time 2) in facial variability and, to a lesser extent, lability, were consistent with Hypothesis 1 resulting in the statistical significance of Contrast 1 (LSVT versus all other groups) and, in the case of facial variability, Contrast 2 (ARTIC versus LSVT). However, neither contrast was significant in the analyses of the complexity of facial expression. To minimize the possible redundancy that may result from the high correlation between AU Variability Δ and AU Lability Δ, the three dependent variables were combined into a canonical

variate (CV) via a discriminant function analysis (DFA), as described next. Although AU Complexity Δ was not highly correlated with other dependent variables (see Table 6), it was included in the DFA for reasons stated in footnote 21.

Step 2: Discriminant function analysis. A DFA with Group as a between-subjects factor and AU Variability Δ, AU Lability Δ, and AU Complexity Δ as dependent variables was carried out to find CVs that provided the best discrimination among the four participant groups. Only the first CV, denoted henceforth "CV_H1," was significant and had the following coefficients:

CV_H1 = 0.78*(AU Variability Δ) + 0.37*(AU Lability Δ) - 0.11*(AU Complexity Δ).

Similar to the preceding analyses, the high equation coefficients of AU Variability Δ and AU Lability Δ suggest that changes in facial variability and, to a lesser extent, lability discriminate best among the four groups. It appears that increases in these variables were observed in LSVT patients, but not other participants (see Figures 7-8). In contrast, changes in the complexity of facial movement did not differentiate the LSVT patients from the other groups.

Step 3: Comparisons of CV_H1 group means. WJ procedures[25] were performed to contrast the LSVT group mean versus the mean of three other groups combined (i.e., Contrast 1) and of the LSVT group mean versus the ARTIC group mean (i.e., Contrast 2). A WJ procedure

---

25 Nonparametric WJ procedures were used for intergroup comparisons because the score distributions compared in these tests violated the assumptions underlying the use of parametric tests. Specifically, Box's M test, $F(18, 7198) = 1.09$, $p > .10$, showed that the homogeneity of variance assumptions was satisfied. However, the assumption of normality was violated: the CV_H1 distribution was negatively skewed and leptokurtic in the ARTIC ($\gamma 1 = -1.67$, $\gamma 2 = 3.77$) and Untreated groups ($\gamma 1 = -1.19$, $\gamma 2 = 2.80$).

showed that the mean CV_H1 score of the LSVT group (M = 0.68, SD = 1.08) was significantly higher than the mean CV_H1 value of the three other groups (M = -0.28, SD = 0.94), F(1, 25.52) = 9.20, p = .01, ES = 0.87. Moreover, the mean CV_H1 value of the LSVT group was significantly higher than that of the ARTIC group (M = -0.21, SD = 1.09), F(1, 23.79) = 4.62, p = .04, ES = 0.82. These results strongly support Hypothesis 1, suggesting that the LSVT patients, relative to the ARTIC and other participants, showed a greater increase on the linear combination of facial movement variability and lability.

Step 4: Eliminating the influence of extraneous variables on intergroup differences on CV_H1. To determine whether demographic and other relevant variables accounted for the differences between LSVT patients and the other three groups on CV_H1, analyses of covariance were undertaken. Specifically, Contrast 1 (LSVT versus all other groups) was tested for statistical significance after the effects of the following covariates were removed: gender, age, handedness, Hoehn-Yahr stage, time since PD diagnosis, education level, BDI score at Time1, MMSE score, and number of PD medications taken.[26] Because of the violation of the normality assumption noted earlier (see footnote 25), a procedure suggested by Hettmansperger (1984, pp. 251-275), a nonparametric analysis of covariance, was carried out separately for each covariate. Unlike other nonparametric alternatives, Hettmansperger's procedure has been shown to

---

26 Because most PD patients took more than one PD medication, it was not possible to compare patients on one medication versus another. Neither was it possible to adequately examine the effect of race as a covariate because 89% of participants were Caucasian (see Table 3). Also, because the Hoehn-Yahr stage, number of PD medications taken, and time since PD diagnosis are characteristics specific to the PD patients, all analyses of the present study tested the effects of these variables based on the data from the PD sample only.

maintain adequate control of Type I and II error rates across a wide range of violations of statistical assumptions (Harwell & Serlin, 1988). Contrast 1 remained significant (all p-values < .05) after the effect of each covariate on CV_H1 was removed, suggesting that the examined variables did not account for the higher CV_H1 mean of LSVT patients relative to that of all other groups combined.

Next, a Hettmansperger's procedure was repeated for Contrast 2 (ARTIC versus LSVT), separately for each covariate. With the effect of gender removed, Contrast 2 became nonsignificant, p = .12. Covarying for other extraneous variables resulted in Contrast 2 becoming marginally significant, all p-values < .10 (see Table 8). Given that the p-value of Contrast 2 in the main analyses was .04, removing the effects of all covariates produced little change in p-values suggesting that the examined variables did not account for the higher CV_H1 mean of LSVT, relative to ARTIC, patients. To summarize, none of the extraneous variables explained the significant differences between CV_H1 group means shown in Contrasts 1 and 2.

Hypothesis 1: Exploratory analyses.

Summary of exploratory analyses. The exploratory analyses of Hypothesis 1 (i.e., hypothesis stating that LSVT, relative to ARTIC and no treatment, reduces facial mobility deficits) were motivated by three aims. The first aim was to further test Hypothesis 1 by performing additional intergroup comparisons. For instance, Hypothesis 1 predicted that changes (from Time 1 to Time 2) in the facial mobility of the LSVT patients would be greater than those of the untreated PD patients. Because the main analyses of Hypothesis 1 were limited to comparing the groups via Contrasts 1 and 2, this prediction was not tested directly in the main analyses. Testing this prediction on each facial mobility measure provided additional evidence regarding which facial mobility aspects change as a result of the LSVT. In addition, to rule out

the possibility that permanent intergroup differences in facial mobility confounded the tests of Hypothesis 1, the relevant intergroup comparisons were performed again, with AU Variability Δ, AU Lability Δ, and AU Complexity Δ measured as a proportion of the participant's overall level of facial mobility rather than in absolute terms (see below).

The second aim was to examine whether the magnitude of facial movement changes in the LSVT group depended on the characteristics of a patient. Although the extraneous variables did not account for the significance of Contrasts 1 and 2, it is possible that the LSVT had a greater effect on a certain patient segment (e.g., patients at a more advanced Hoehn-Yahr stage).

Finally, the third aim of the exploratory analyses was to replicate the findings of the DFA (discriminant function analysis) in a more homogenous sample that consisted entirely of PD patients. Such replication would provide additional support that the high coefficients of AU Variability Δ and AU Lability Δ in the CV_H1 formula reflected LSVT-related changes in facial variability and lability rather than the idiosyncrasies of our participant sample.

As reported next, the additional tests of Hypothesis 1 were broadly consistent with the results of the main analyses. The exploratory analyses also showed that the effect of LSVT tended to be greater for patients with a higher education level. Further, the results of the exploratory DFA paralleled the findings of the DFA in the main analyses.

Additional intergroup comparisons to test Hypothesis 1. To more comprehensively examine Hypothesis 1 and possibly reveal other important findings, two sets of analyses were conducted. In the first set, additional contrasts of CV_H1, AU Lability Δ, and AU Variability Δ group means were carried out. In the second set, the same intergroup contrasts were repeated on relative rather than absolute changes in facial mobility.

Additional comparisons of group means on dependent variables. Because neither Contrast 1 (LSVT versus all other groups) nor Contrast 2 (ARTIC versus LSVT) of AU Complexity Δ group means was significant (see the "Main Analyses" subsection), no additional comparisons of the means on this variable were performed. For other dependent variables (CV_H1, AU Lability Δ, and AU Variability Δ), WJ procedures were performed to compare the group means, separately on each dependent variable, via the following contrasts: LSVT versus Untreated, LSVT versus Control, and ARTIC versus Untreated.

For CV_H1, the mean of the LSVT group (M =0.68, SD = 1.08) was significantly greater than that of the Untreated (M = -0.30, SD = 0.87), $F(1, 28.78) = 8.22$, $p = .01$, ES = 0.90, and Control groups (M = -0.30, SD = 0.95), $F(1, 23.38) = 6.17$, $p = .02$, ES = 0.90. Unlike the LSVT versus Untreated contrast, the CV_H1 means of the ARTIC (M = -0.21, SD = 1.09) and Untreated groups were not significantly different, $F(1, 20.31) = 8.22$, $p = .81$.

For AU Variability Δ, the mean of the LSVT group (M = 1.81, SD = 3.04) was significantly greater than that of the Untreated (M = -0.94, SD = 2.33), $F(1, 28.13) = 8.46$, $p < .01$, ES = 0.88, and Control groups (M = -0.73, SD = 2.65), $F(1, 23.46) = 5.31$, $p = .03$, ES = 0.81. Unlike the LSVT versus Untreated contrast, the AU Variability Δ means of ARTIC and Untreated groups were not significantly different, $F(1, 19.19) = 0.11$, $p = .74$.

For AU Lability Δ, the mean of the LSVT patients (M = 5.00, SD = 8.23) was significantly greater than that of the Untreated group (M = -1.71, SD = 6.82), $F(1, 30.95) = 7.34$, $p = .01$, ES = 1.05, but not that of the Control group (M = -1.18, SD = 7.77), $F(1, 20.96) = 1.44$, $p = .24$). In contrast to the LSVT versus Untreated comparison, the AU Lability Δ means of the ARTIC and Untreated groups were not significantly different, $F(1, 25.42) = 1.37$, $p = .25$.

To sum up, the facial mobility changes observed in the LSVT group were greater than the corresponding changes in the Untreated group on all of the variables examined in these additional intergroup comparisons. In contrast, the facial mobility changes of the ARTIC and Untreated groups were not significantly different on any variable. Awaiting replication in further studies, these preliminary results are in line with Hypothesis 1.

Testing Hypothesis 1 with the relative measures of change in facial mobility. Because the analyses in the preceding subsections are exploratory, these results need to be interpreted with caution. To provide a stronger basis for interpretation, relevant intergroup contrasts were repeated on relative changes in facial mobility (i.e., in contrast to the absolute change levels used so far). Namely, a relative change in AU Lability, denoted AU Lability Δ%, was calculated according to the following formula:

AU Lability Δ% = (AU Lability at Time 2 - AU Lability at Time 1)/(AU Lability at Time 1 + AU Lability at Time 2).

Relative changes on two other facial mobility measures, AU Variability Δ% and AU Complexity Δ%, were calculated in the same manner. Using this method, changes in participants' facial mobility are measured relative to their facial mobility level across Times 1 and 2. This approach lessens the effect of interindividual differences on the amount of facial mobility on the tests of Hypothesis 1.

Next, WJ procedures[27] were performed to contrast intergroup means on these dependent measures. These procedures revealed that the AU Complexity Δ% mean of the LSVT group (M

27 Nonparametric WJ procedures were chosen for intergroup comparisons because the score distributions compared in these tests violated the assumptions underlying the use of parametric tests. Though the modified O'Brien's procedures showed that for all three measures,

= 0.14, SD = 0.36) was not significantly different from that of the other groups combined (M = 0.04, SD = 0.28), F(1, 23.39) = 0.94, p = .34, or that of the ARTIC group (M = 0.05, SD = 0.45), F(1, 20.65) = 0.40, p = .53. Because these main contrasts (Contrasts 1 and 2) did not reveal any statistically significant differences, no additional analyses of AU Complexity Δ% scores were performed.

For AU Variability Δ%, the mean of the LSVT group (M = 0.23, SD = 0.39) was significantly greater than that of the Untreated (M = -0.09, SD = 0.30), F(1, 28.40) = 6.87, p = .01, ES = 0.61, and that of the other groups combined (M = -0.04, SD = 0.37), F(1, 26.56) = 5.01, p = .03, ES = 0.49, but not significantly different from the mean of the ARTIC (M = 0.02, SD = 0.53) group, F(1, 19.50) = 1.41, p > .10. Unlike the LSVT versus Untreated contrast, the AU Variability Δ% means of the ARTIC and Untreated groups were not significantly different, F(1, 16.20) = 0.40, p > .10. The difference between the AU Variability Δ% means of the LSVT and Control (M = -0.01, SD = 0.25) groups was marginally significant, F(1, 24.96) = 3.76, p = .06, ES = 0.46.

---

the score distributions within each participant group did not violate the homogeneity of variance assumption, skewness and kurtosis values showed that the assumption of normality was violated. Specifically, in the AU Complexity Δ% distributions, there was positive kurtosis in the ARTIC ($\gamma 2 = 1.65$) and Control ($\gamma 2 = 2.44$) groups, positive skew in the Control group ($\gamma 1 = 1.69$), and negative skew in the ARTIC group ($\gamma 1 = -1.16$). In the AU Lability Δ% distributions, there was positive kurtosis in the ARTIC ($\gamma 2 = 3.92$) and LSVT ($\gamma 2 = 11.70$) groups *and* positive skew in the ARTIC ($\gamma 1 = 1.80$), LSVT ($\gamma 1 = 3.32$), and Untreated groups ($\gamma 1 = 1.27$). In the AU Variability Δ% distributions, there were positive skew and kurtosis in the ARTIC ($\gamma 1 = 1.69$, $\gamma 2 = 3.67$), LSVT ($\gamma 1 = 1.33$, $\gamma 2 = 3.01$), and Control ($\gamma 1 = 2.01$, $\gamma 2 = 5.01$) groups.

For AU Lability Δ%, the mean of the LSVT patients (M = 0.27, SD = 0.44) was significantly greater than that of the Untreated group (M = -0.08, SD = 0.30), $F(1, 26.43) = 6.84$, $p = .01$, ES = 0.59, or that of the other groups combined (M = -0.04, SD = 0.38), $F(1, 24.97) = 5.42$, $p = .03$, ES = 0.51. The AU Lability Δ% mean of the LSVT group was not significantly different from that of the ARTIC group (M = -0.01, SD = 0.58), $F(1, 19.79) = 1.96$, $p > .10$. In contrast to the LSVT versus Untreated comparison, the AU Lability Δ% means of the ARTIC and Untreated groups were not significantly different, $F(1, 15.23) = 0.13$, $p = .73$. The difference between the AU Lability Δ% means in the LSVT and Control (M = 0.00, SD = 0.23) groups was marginally significant, $F(1, 23.53) = 4.22$, $p = .05$, ES = 0.46.

To summarize, the relative measures of facial mobility changes yielded results that were broadly consistent with those based on the absolute measures. As with the absolute measures, Contrast 1 (LSVT versus all other groups) was significant in the analysis of facial lability and variability. Contrary to the predictions of Hypothesis 1, the direct comparison of LSVT versus ARTIC means was not significant with any relative measure. However, similar to the results based on the absolute measures, the increases in the facial lability and variability of the LSVT, but not ARTIC, patients were significantly greater than the changes in the Untreated group. Taken together, the findings based on the absolute and relative measures of changes in facial mobility moderately support Hypothesis 1 and are in line with the results of the main analyses.

The role of a patient's characteristics in the effect of LSVT on facial movement.

Although the extraneous variables did not account for the significance of Contrasts 1 and 2, it is possible that these variables moderated the effect of LSVT revealed in these contrasts (in other words, the magnitude of differences between CV_H1 means across groups depended on these variables). For example, facial mobility may have improved to a greater extent in older than

younger LSVT patients. To investigate the possibility of statistical moderation, the effect of extraneous variables on the dependent variables (CV_H1, AU Lability Δ, or AU Variability Δ) was tested using the data from the LSVT group only. The following extraneous variables were examined: gender, age, Hoehn-Yahr stage, time since PD diagnosis, education level, BDI score at Time 1, BDI score at Time 1 dichotomized into "low" (BDI score ≤ 13) versus "moderate" (BDI score > 13) categories,[28] MMSE score, and number of PD medications taken.[29]

When the normality or homogeneity of variance assumptions were met,[30] parametric procedures were carried out. Specifically, for discrete extraneous variables (gender, handedness,

---

28 The cutoff point of 13/14 was chosen because a 0-13 point range on the 2nd version of the Beck Depression Inventory (BDI-II; Beck, 1967) is generally considered as indicative of "minimal depression" (Beck et al., 1996). This cutoff resulted in 11 LSVT patients being classified as having low levels of depressive symptomatology and 5 patients as having mild levels of depressive symptomatology. As noted earlier, individuals possibly having severe depression (scores above 24 on the BDI-II) were not recruited in the present study (see the Method section).

29 As there was only one left-handed LSVT patient, the effect of handedness on LSVT-related improvements in facial mobility could not be explored.

30 Except the CV_H1 and Variability Δ score distributions of male LSVT patients, which had a slight positive skew ($\gamma 1 = 1.03$ and $\gamma 1 = 1.11$, respectively), all other score distributions examined in these tests did not violate the normality or homogeneity of variances assumptions, as shown by modified O'Brien's procedures and the examination of outliers, stem and leaf plots, box and whiskers plots, and absolute skewness and kurtosis values, none of which exceeded 1.

and the dichotomized BDI score), a parametric analysis of variance (ANOVA) with an extraneous variable as a between-subjects factor was carried out separately for each dependent variable. For continuous extraneous variables (age, Hoehn-Yahr stage, time since PD diagnosis, education, BDI score at Time 1, MMSE score, and number of PD medications taken by the patient), a Pearson correlation was calculated for each combination of an extraneous variable and dependent variable.

Participant education level was significantly correlated with CV_H1, $r = .59$, $n = 16$, $\underline{p} = .02$, AU Lability Δ, $r = .56$, $n = 16$, $\underline{p} = .02$, and AU Variability Δ, $r = .59$, $n = 16$, $\underline{p} = .02$. Possibly due to the small sample size of 16 LSVT patients, no other effects were significant.[31]

Due to the violation of the normality assumption, WJ procedures were performed to examine the effect of gender on the CV_H1 and AU Variability Δ scores of LSVT patients.

31 Based on the evidence of reduced facial expressivity in depression (Gaebel & Wölwer, 2004; Jaeger, Borod, & Peselow, 1986) and the comorbidity of PD and depression (Burn, 2002), we explored whether PD patients' depressive symptoms were associated with reduced facial mobility by computing bivariate Pearson correlations between PD patients' BDI scores at Time 1 and their values on each facial movement variable (AU Variability, AU Lability, and AU Complexity) at Time 1 within the entire sample of PD patients. This sample included the ARTIC, LSVT, and Untreated groups whose mean BDI scores did not significantly differ (see the Method section and Table 3). Bivariate Pearson correlations between the BDI scores of patients in the ARTIC and Untreated groups and their values on each facial movement variable at Time 2 were also determined (to avoid confounding the influence of depression with the effects of voice treatment, the LSVT patients were excluded from this analysis). No significant effects were obtained, all $\underline{p}$-values $> .10$. Associations between dichotomized BDI

Auxiliary DFA (discriminant function analysis) of PD patients' changes on facial mobility variables. The coefficients of the CV_H1 equation suggest that LSVT-related improvement in facial mobility is characterized primarily by increases in facial variability and lability. To confirm that these findings were not produced as an artifact of the differences between the facial movements of PD patients and controls, the DFA with Group as a between-subjects factor and AU Variability Δ, AU Lability Δ, and AU Complexity Δ as dependent variables was performed after excluding Control group data (i.e., only data from PD patients were included). Possibly due to small sample size, no CVs reached statistical significance (all p-values > .10), and, as a result, one cannot confidently interpret the results of this exploratory DFA. However, it is noteworthy that the equation coefficients of the first CV, termed CV_H1a, were similar to those of CV_H1:

CV_H1a = 0.62*(AU Variability Δ) + 0.39*(AU Lability Δ) + 0.07*(AU Complexity Δ).

Supporting the results of the main DFA, these equation coefficients suggest that, independent of the differences between the facial movements of PD patients versus age-matched controls, LSVT-related improvement in facial mobility is characterized by increases in facial variability and, to a lesser degree, lability rather than by changes in the complexity of facial movement.

Hypotheses 2a and 2b

---

scores and the values of each facial movement variable (both at Time 1 and Time 2) were explored in a similar fashion, except that a WJ procedure or ANOVA (i.e., depending on whether the assumptions underlying parametric tests were violated) instead of a Pearson correlation was conducted for each comparison. No significant effects were obtained, all p-values > .10.

Overview of analyses. We made the following hypotheses regarding the perception of PD patients by others. Hypothesis 2a stated that this study would replicate the negative impressions of PD patients obtained in earlier studies. Hypothesis 2b posits that the LSVT, relative to the articulation treatment or no treatment, reduces these negative judgments.

To avoid reducing statistical power of the main analyses by making a high number of intergroup comparisons on each personality and behavior scale, each observer's raw ratings of participants from a given group (e.g., ARTIC) were reduced to three composite ratings of personality, behavior, and overall perception of a patient (i.e., termed Pers+, Beh+, and Overall+, respectively), as described next. The dependent variables in the main analyses were limited to these three measures. In contrast, the exploratory analyses were carried out with both composite and raw ratings (see the "Exploratory Analyses of Hypotheses 2a and 2b" subsection below). Also, analogous to the analyses of Hypothesis 1, the main analyses were limited to only planned comparisons (e.g., Contrasts 1 and 2), whereas the exploratory analyses included other relevant comparisons.

Calculation of the composite ratings. Overall+, Pers+, and Beh+ ratings that were used as three dependent variables in the main analyses of Hypotheses 2a and 2b were calculated as follows. First, because there were seven points on all scales, each raw rating was transformed into a number ranging from 1 to 7, with 1 and 7 corresponding to the negative and positive ends of a scale, respectively. Then, to calculate the Pers+ rating of the ARTIC group at Time 1 by a given observer, the Time 1 ratings on all personality scales (introverted - extraverted, secure - anxious, independent - dependent, and disliked - liked) were averaged across all ARTIC participants viewed by that observer. Thus, the resulting Time 1 Pers+ rating was a measure of how positively, on average, a given observer judged the personality of an ARTIC participant at

Time 1. Next, the same procedure was repeated with the ratings of the LSVT, Control, and Untreated participants, separately for each group. As a result, 440 Time 1 Pers+ ratings (110 observers x 4 groups) in total were derived.

Identical procedures were performed with the ratings on the behavior scales (tense - relaxed, sad - happy, engaged - bored, openhearted - guarded, and disliked - liked) to calculate the mean Time 1 Beh+ ratings of each group. The resulting Time 1 Beh+ ratings were a measure of how positively, on average, a given observer judged the Time 1 behavior of a participant from a given group (e.g., an ARTIC participant). Finally, each observer's Overall+ ratings of a given participant group were calculated by averaging his or her Pers+ and Beh+ ratings of that group. The resulting Time 1 Overall+ ratings were a measure of how positively, on average, an observer judged the personality and behavior of a participant from a given group at Time 1. Time 2 Overall+, Pers+, and Beh+ ratings were calculated using the same method as the Time 1 ratings. Mean Overall+, Pers+, and Beh+ ratings of each participant group at Times 1 and 2 are shown in Figures 10-12 and Table 9. This approach of analyzing the ratings averaged across participants rather than observers (i.e., using observers instead of participants as the unit of analysis) afforded more statistical power because the number of observers (n = 110) substantially exceeded the number of participants (n = 56).

Hypothesis 2a: Main analyses. Three dependent variables were used in the main analyses of Hypothesis 2a (i.e., the hypothesis stating that the negative perceptions of PD patients that were obtained in earlier studies would be replicated in this study): Overall+, Pers+, and Beh+ ratings. These analyses involved two steps. First, the Time 1 Pers+, Beh, and Overall+ ratings of PD patients were compared with those of controls, separately on each dependent variable. Second, analyses of covariance were carried out to examine whether

eliminating the influence of extraneous variables affects the outcomes of intergroup comparisons. As described next, the results supported Hypothesis 2a showing that PD patients were rated more negatively than age-matched controls on each dependent variable and that none of the extraneous variables accounted for these differences between the two groups.

Statistical comparisons of PD patients versus controls on the composite ratings. To replicate the findings of earlier studies (e.g., Pentland et al., 1988) showing the perceived differences between the personalities of PD patients and those of controls (i.e., test Hypothesis 2a), a WJ procedure[32] was performed to contrast the 330 Time 1 Pers+ ratings of PD participants (110 observers x 3 groups) with 110 Time 1 Pers+ ratings of controls. The same WJ procedures were carried out separately on the Overall+ and Beh+ ratings.[33] In addition to these contrasts,

---

32 Nonparametric WJ procedures were selected for testing Hypothesis 2a because the score distributions compared in these tests violated the homogeneity of variance assumption. Although the distributions of Overall+, Beh+, and Pers+ ratings of both PD patients and controls appeared to satisfy the assumption of normality, the modified O'Brien procedure showed that the variance of ratings was not homogenous across the two groups (PD patients and controls), $F(1, 438) = 38.52$, $p < .05$ for Overall+, $F(1, 438) = 29.66$, $p < .05$ for Pers+, and $F(1, 432) = 16.53$, $p < .05$ for Beh+.

33 Statistical analyses were carried out to examine whether the Time 1 means of the PD groups on each dependent variable were similar. Given the violations of the homogeneity of variance assumption in the Overall+ and Pers+ score distributions, a WJ procedure was carried out for this purpose, separately for each of these two variables. Because the Beh+ score distributions met the assumptions underlying parametric tests, a one-way ANOVA with Group as an independent variable was performed on these ratings. For each dependent variable, the effect

other relevant statistical comparisons of the group means were performed to more comprehensively examine the dataset. However, because these additional comparisons were post hoc, their results are reported in the "Exploratory Analyses" subsection below.

In support of Hypothesis 2a, the Overall+, Pers+, and Beh+ mean ratings of PD participants were significantly lower than those of controls, $\underline{F}(1, 141.21) = 42.87$, $\underline{p} < .01$, ES = 0.90 for Overall+, $\underline{F}(1, 143.60) = 21.62$, $\underline{p} < .01$, ES = 0.65 for Pers+, $\underline{F}(1, 147.19) = 57.46$, $\underline{p} < .01$, ES = 0.97 for Beh+. These results show that the personality and behavior of PD participants were perceived more negatively than the personality and behavior of controls.

---

of Group was significant, $F(2, 214.41) = 20.87$, $\underline{p} < .01$, ES = 0.19 for Overall+, $F(2, 214.63) = 10.63$, $\underline{p} < .01$, ES = 0.15 for Pers+, and $F(2, 322) = 16.34$, $\underline{p} < .01$, ES = 0.09 for Beh+. Post hoc comparisons showed that all types of ratings of the LSVT group were significantly lower than the corresponding ratings of the Untreated or ARTIC groups, that all $\underline{p}$-values < .01, and that the Beh+ ratings of the ARTIC group were significantly lower than the corresponding ratings of the Untreated group, $\underline{p} < .05$.

The effect sizes of the obtained differences, ranging from 0.09 to 0.19, were, on average, roughly six times smaller than the effect sizes of the corresponding differences between PD patients and controls, ranging from 0.65 to 0.97. This dissimilarity in magnitude suggests that the two types of differences -- the ones obtained among PD patient groups versus the one found between PD patients and controls -- arise for different reasons. Possibly reflecting normal variation in how individuals are perceived, the differences among PD patient groups might have occurred because the PD patient groups were not matched on their perceived personality and behavior characteristics during recruitment.

Eliminating the effect of extraneous variables on the intergroup comparisons of Pers+, Beh+, and Overall+ ratings. The means contrasted with one another to test Hypothesis 2a were based on the ratings made by different groups of observers (see the Method section). Although, as shown in the Method section, the demographic characteristics did not significantly differ across the observer groups, it is possible that the influence of these characteristics on observers' ratings accounts for some of the findings supporting Hypothesis 2a. Other extraneous variables, such as Order (i.e., whether an observer made personality versus behavior judgments first), may have similar effects.

To rule out such confounding effects, a Hettmansperger's procedure was performed to determine whether the results obtained in the main analyses remained the same after the influence of extraneous variables on the ratings was removed. As discussed earlier, Hettmansperger's procedure, a nonparametric analysis of covariance, removes the influence of an extraneous variable, entered into an analysis as a covariate, on a dependent variable. The following covariates were examined: Order, Set (i.e., which of the four sets of clips was shown to an observer), Sequence (i.e., which of the two sequences of a given set of clips was shown to the observer), observer's gender, observer's age, observer's race, and observer's education level. In all analyses described in this section, a separate Hettmansperger's procedure was carried out for every possible combination of a covariate with a dependent variable (Pers+, Beh+, or Overall+ ratings). For each combination of a covariate and dependent variable, the Time 1 ratings of PD patients remained significantly lower than those of controls (all p-values $< .01$). Thus, the examined demographic variables did not account for the findings supporting Hypothesis 2a.

Hypothesis 2b: Main analyses. To examine whether, as Hypothesis 2b states, LSVT, relative to the articulation treatment or no treatment, reduces the negative judgments of PD patients by others, changes (from Time 1 to Time 2) in observers' ratings of LSVT patients were compared with the ratings of other patients. Namely, changes in the ratings of the LSVT group were contrasted with those of the non-LSVT groups (Contrast 1) and those of the ARTIC group (Contrast 2), separately for each of the three rating types (Pers+, Beh+, and Overall+). Next, analyses of covariance were conducted to determine if eliminating the effect of an extraneous variable affected the results yielded by these contrasts. As detailed next, the results of these main analyses provided little support for Hypothesis 2b showing that the amount or direction of change in the perception of participants by others did not differ significantly across groups. Moreover, these results remained the same when the effects of the extraneous variables were taken into account.

Planned contrasts of the Overall+, Pers+, and Beh+ ratings. A WJ procedure[34] was performed to compare the group means via Contrasts 1 and 2, separately on each dependent variable. The results of these comparisons are reported next. Besides Contrasts 1 and 2, other

---

34 Nonparametric WJ procedures were chosen for testing Hypothesis 2b because both the homogeneity of variance and normality assumptions were violated in the distributions of scores on all three variables. Specifically, there was positive kurtosis in the distributions of the Overall+ ($\gamma 2 = 1.86$) and Pers+ ($\gamma 2 = 1.50$) ratings of the ARTIC group at Time 2. Also, the distribution of Beh+ ratings of the LSVT group at Time 2 was negatively skewed and leptokurtic ($\gamma 1 = -1.40$, $\gamma 2 = 4.42$). The modified O'Brien procedure showed that for each variable, the variance of ratings was not homogenous across the four participant groups, $F(7, 872) = 13.04$, $p < .05$ for Overall+, $F(7, 872) = 10.74$, $p < .05$ for Pers+, and $F(7, 859) = 7.58$, $p < .05$ for Beh+.

relevant statistical comparisons of the group means were performed to more comprehensively examine the dataset. However, because these additional comparisons were post hoc, their results are reported in the "Exploratory Analyses" subsection below.

WJ procedures revealed that neither contrast was significant for any of the three variables (all p-values > .10). It should be noted, however, that Contrast 2 (ARTIC versus LSVT) was marginally significant for Pers+ ratings, $F(1, 414.89) = 3.07$, $p = .08$, ES = 0.29. This result reflects that the Time 2 ratings of the LSVT group (M = 4.08) were greater than the Time 1 ratings of the same group (M = 3.87), whereas the ratings of the ARTIC group at Times 1 and 2 were nearly the same (M = 4.17 for Time 1 and M = 4.15 for Time 2). Thus, the results of the main analyses did not provide strong support for Hypothesis 2b. The means compared in Contrasts 1 and 2 and the corresponding statistics are shown in Table 10.

Eliminating the effect of extraneous variables on the intergroup comparisons of Pers+, Beh+, and Overall+ ratings. Similar to the analyses of Hypothesis 2a, the tests of Hypothesis 2b may have been confounded by the effect of extraneous variables. To eliminate these confounding effects, analyses of covariance were performed. Specifically, a separate Hettmansperger's procedure was performed for every possible combination of a covariate with a dependent variable (Pers+, Beh+, or Overall+ ratings). The same covariates as in the analyses of Hypothesis 2a were examined: Order, Set (i.e., which of the four sets of clips was shown to an observer), Sequence (i.e., which of the two sequences of a given set of clips was shown to the observer), observer's gender, observer's age, observer's race, and observer's education level. As previously, each extraneous variable was entered into an analysis as a covariate.

The results of both contrasts remained essentially the same. Specifically, Contrast 2 (ARTIC versus LSVT) remained marginally significant (i.e., the p-values ranged from .04 to .06,

depending on a covariate) when the Pers+ ratings were analyzed. No other significant findings were revealed for either contrast (all p-values > .10).

Exploratory analyses of Hypotheses 2a and 2b.

Summary of exploratory analyses. Four sets of exploratory analyses of Hypotheses 2a and 2b were conducted. The first set explored which aspects of personality and behavior were perceived to be more negative in PD patients than in others. These analyses aimed to provide a more detailed description of the negative impressions, revealed in the tests of Hypothesis 2a, of PD patients. To this end, the raw ratings of PD patients were contrasted with those of controls on each personality and behavior scale.

The second set of analyses examined more thoroughly whether variable Time (Time 1 versus Time 2) had an effect on the ratings of each participant group. Because Contrasts 1 and 2 were targeted to identify specific Time x Group interactions, they do not detect the within-group effects of variable Time. Such effects may qualify the lack of significant effects in the main tests of Hypothesis 2b (i.e., hypothesis stating that the LSVT improves the negative perceptions of PD patients) and have implications for future studies. To reveal such effects, additional comparisons that were not included in the main analyses were carried out.

The third and fourth sets of analyses investigated whether observer and participant characteristics, respectively, played a role in the negative judgments of PD patients. To address this issue, the effect of such characteristics (e.g., gender and time since PD diagnosis) on Time 1 ratings was examined. The final set of analyses explored whether the magnitude of the LSVT's effect depended on patient characteristics. To investigate this question, the effect of these characteristics on the amount of change (from Time 1 to Time 2) in the ratings of the LSVT group was probed.

To summarize, these analyses showed that PD patients were rated more negatively than controls on nearly all personality and behavior scales. Moreover, additional comparisons provided preliminary support for the hypothesis that the LSVT partly alleviates the negative impressions of PD patients. Furthermore, female observers perceived PD patients more positively than male observers, and PD patients with more depressive symptoms tended to be perceived more negatively and show smaller changes (from Time 1 to Time 2) in the Pers+ and Overall+ ratings than did patients with fewer symptoms.

Characterizing the negative perception of PD patients by others. To determine which personality and behavioral aspects were perceived to be more negative in PD patients than controls, WJ procedures[16] were conducted to contrast the Time 1 ratings of controls with those of the other three groups combined, separately for each personality and behavior scale. As the "Liked versus Disliked" scale appeared on both personality and behavior forms (i.e., as noted in the Method section, observers rated how much they liked each participant twice), the average rating across these identical scales was entered in the analysis.

Replicating and extending prior findings, the analysis showed that PD patients were perceived to be more introverted, $F(1, 142.97) = 49.18$, $p < .01$, ES = 0.88, anxious, $F(1, 149.48) = 9.74$, $p < .01$, ES = 0.46, bored, $F(1, 176.28) = 38.05$, $p < .01$, ES = 0.63, tense, $F(1, 154.66) = 39.89$, $p < .01$, ES = 0.81, sad, $F(1, 152.71) = 40.19$, $p < .01$, ES = 0.82, guarded, $F(1, 154.66) = 39.89$, $p < .01$, ES = 0.66, and less liked, $F(1, 160.73) = 36.08$, $p < .01$, ES = 0.71 than controls. In contrast, PD participants were not judged to be significantly more or less dependent than controls, $p = .41$. The means of both groups on each scale and the corresponding statistics are presented in Table 11 and Figure 13.

Within-group effects of Time on the ratings of personality and behavior. To clarify whether observers' perceptions of each participant group at Time 1 versus Time 2 were different, a number of exploratory analyses were carried out. In the first set of analyses, the Time 1 versus Time 2 ratings (see Figures 10-12 and Table 9) were contrasted within each group by performing WJ procedures. The findings revealed by the analysis of the Overall+ ratings are presented first. For the Control group, an increase from Time 1 (M = 4.74, SD = 0.93) to Time 2 (M = 5.08, SD = 0.70) was significant, F(1, 213.40) = 3.77, p = .01, ES = 0.46. For the LSVT group, an increase from Time 1 (M = 3.85, SD = 0.53) to Time 2 (M = 4.03, SD = 0.66) was also significant, F(1, 209.35) = 5.01, p = .03, ES = 0.25. In contrast, the same comparisons did not reach significance in the ARTIC or Untreated groups (all p-values > .10).

For the Pers+ ratings, similar results, discussed next, were obtained. For the Control group, the rating increase from Time 1 (M = 4.55, SD = 0.97) to Time 2 (M = 4.97, SD = 0.93) was significant, F(1, 217.47) = 10.62, p < .01, ES = 0.53. In the LSVT group, the increase from Time 1 (M = 3.87, SD = 0.57) to Time 2 (M = 4.08, SD = 0.68) was also significant, F(1, 210.54) = 5.73, p = .02, ES = 0.26. Changes from Time 1 to Time 2 were not significant in the ARTIC or Untreated groups (all p-values > .10).

Finally, with respect to the Beh+ ratings of the Control group, the increase from Time 1 (M = 4.94, SD = 1.02) to Time 2 (M = 5.20, SD = 0.97) was marginally significant, F(1, 213.40) = 3.77, p = .05, ES = 0.30. Changes in the Beh+ ratings were not significant in any other groups (all p-values > .10).

These preliminary findings suggest that controls at Time 2 were judged more positively than the same controls at Time 1 in terms of their personality. Because Hypothesis 2b (i.e., hypothesis stating that the LSVT improves the negative judgments of PD patients) does not

predict that the perception of any participants other than LSVT patients would improve, it can account for only some of the observed results. On the other hand, because the LSVT group, unlike the ARTIC or Untreated patients, was perceived more positively after than before treatment, it is possible that, in line with Hypothesis 2b, LSVT may slightly improve the perception of PD patients. Although, as suggested by the obtained effect sizes of 0.25-0.26, this improvement is modest, the effect of the LSVT may be uneven and manifest in the perception of some personality aspects more than in others. Thus, to probe a significant change of the Pers+ ratings observed in the LSVT group, the Time 1 ratings of the LSVT group were contrasted with the Time 2 ratings by performing WJ procedures, separately for each personality scale. Results indicated that LSVT patients were judged to be more extraverted, $F(1, 208.59) = 13.41$, $p < .01$, ES = 0.37, independent, $F(1, 208.26) = 4.43$, $p = .04$, ES = 0.24, secure, $F(1, 214.26) = 12.07$, $p < .01$, ES = 0.46, and liked, $F(1, 217.99) = 13.51$, $p < .01$, ES = 0.42, after than before treatment. In sum, these preliminary results suggest that the LSVT had a positive effect on the perception of several personality aspects in PD.

The role of observer characteristics in the judgments of PD patients' personality and behavior. Although the demographics of observers did not fully account for the negative effect of PD on their judgments, it is possible that these demographics partly contributed to how PD patients were perceived. To clarify whether observer demographics (gender, race, age, and education level) affected observers' perception of PD patients, a separate WJ (for the discrete variables of gender and race) or nonparametric least trimmed squares regression (for the continuous variables of age and education level) procedure was performed for all possible combinations of a demographic variable and rating type (Overall+, Pers+, and Beh+). In each procedure, a demographic variable of interest and PD Status (PD versus control) were entered as

independent variables, and Time 1 ratings of participants were entered as a dependent variable. Nonparametric procedures were chosen because modified O'Brien procedures showed that the variance of ratings across the two groups (PD patients and controls) was not homogenous, regardless of the rating type (all p-values < .05).

There were no significant interactions between PD status and any demographic variables (all p-values > .10). There was a significant main effect of Gender on Overall+, $F(1, 133.13) = 5.91$, $p = .02$, ES = 0.68, with the ratings made by the female observers (M = 4.41, SD = 0.82) being significantly higher than those made by the male observers (M = 4.25, SD = 0.70). Similarly, the Pers+ ratings made by the female observers (M = 4.39, SD = 0.82) were significantly higher than those made by the male observers (M = 4.22, SD = 1.01), $F(1, 138.44) = 9.14$, $p < .01$, ES = 0.80. The effect of gender on the Beh+ ratings was marginally significant, $F(1, 132.14) = 2.76$, $p = .099$, ES = 0.45. No other effects were significant (all p-values > .10).

The role of participant characteristics in the perception of PD patients. Similar to observer characteristics, participant characteristics may have played a role in the negative perceptions of PD patients. Clearly, ratings averaged across participants could not be used for examining the effects of participant demographics on these impressions. Thus, for the analyses in this subsection, the Time 1 raw ratings of each of the 56 participants were averaged across observers, separately for each rating type (Pers+, Beh+, or Overall+). Thus, the resulting dataset consisted of 56 Overall+, 56 Pers+, and 56 Beh+ ratings of participants at Time 1.

The following participant characteristics were examined: gender, age, handedness, Hoehn-Yahr stage, time since PD diagnosis, education level, BDI score at Time 1, MMSE score,

and the number of PD medications taken by the patient. A separate linear regression analysis[35] with a participant characteristic and PD Status as two independent variables and ratings as a dependent variable was performed for every possible combination of a demographic variable and rating type. The effect of gender on Beh+ ratings was marginally significant, $F(1, 52) = 2.86$, $p = .097$, with female participants ($M = 4.72$, $SD = 0.89$) receiving higher ratings than male participants ($M = 4.08$, $SD = 0.65$). Also, the BDI scores at Time 1 had a marginally significant main effect on Overall+, $F(1, 52) = 2.87$, $p = .096$, and Pers+, $F(1, 52) = 2.98$, $p = .09$, ratings. Moreover, there were significant interactions between the BDI scores at Time 1 and PD Status for Overall+, $F(1, 52) = 7.77$, $p < .01$, Pers+, $F(1, 52) = 8.55$, $p < .01$, and Beh+, $F(1, 52) = 6.64$, $p < .01$, ratings. No other effects were significant (all p-values > .10).

To interpret the BDI score x PD Status interactions, Pearson correlations between the BDI scores and each type of rating were computed separately for PD patients and controls. As shown in Table 12 and Figures 14-16a & b, the BDI scores of PD patients were negatively correlated with each of the three ratings, all p-values < .05. In contrast, the BDI scores of controls were positively, albeit not significantly, correlated with all three ratings. Therefore, the personality and behavior of PD patients with higher BDI scores tended to be perceived more

---

35 Parametric linear regression was chosen to conduct this analysis because all rating distributions analyzed in these tests (e.g., the Pers+ ratings of male participants in the tests of gender's effect on Pers+) met the normality and homogeneity of variance assumptions as shown by the results of modified O'Brien's procedures *and* the examination of outliers, stem and leaf plots, box and whiskers plots, and absolute skewness and kurtosis values none of which exceeded 1.

negatively than those of PD patients with lower scores.[36] This relationship did not hold in the Control group.

To determine which personality and behavioral aspects were perceived to be more negative in PD patients with higher BDI scores, Pearson correlations between patients' BDI scores and the Time 1 ratings of these individuals were computed, separately for each personality and behavior scale. As the "Liked versus Disliked" scale appeared on both personality and

---

36 The same relationships between depressive symptoms and ratings were revealed when instead of a raw BDI score, a dichotomized BDI score (i.e., "low" versus "moderate"; see footnote 12) was entered in the analysis. Specifically, a parametric analysis of variance (ANOVA), performed separately for each dependent variable, showed that the Overall+, Pers+, and Beh+ ratings of PD patients at Time 1 with low BDI scores were significantly higher than the ones of PD patients with moderate BDI scores, $F(1, 43) = 5.64$, $p < .05$, $\eta 2 = .12$ for Overall+, $F(1, 43) = 6.03$, $p < .05$, $\eta 2 = .12$ for Pers+, and $F(1, 43) = 4.89$, $p < .05$, $\eta 2 = .10$ for Beh+. Similarly, the Overall+, Pers+, and Beh+ ratings of PD patients at Time 2 with low BDI scores were significantly higher than the ones of PD patients with moderate BDI scores, $F(1, 43) = 5.39$, $p < .05$, $\eta 2 = .11$ for Overall+, $F(1, 43) = 6.18$, $p < .05$, $\eta 2 = .13$ for Pers+, and $F(1, 43) = 4.39$, $p < .05$, $\eta 2 = .09$ for Beh+. The Overall+, Pers+, and Beh+ group means of both groups of PD patients are shown in Table 13.

ANOVAs were chosen for these group comparisons because the tests of skewness and kurtosis showed that the Pers+, Beh+, and Overall+ score distributions of both groups (PD patients with low and moderate BDI scores) met the normality assumption. Also, modified O'Brien's procedures demonstrated that the homogeneity of variance assumption was satisfied in each group comparison.

behavior forms, the average rating across these identical scales was entered in the analysis. The analysis showed that PD patients with higher BDI scores tended to be perceived as more introverted, $r = -.36$, $p < .05$, dependent, $r = -.40$, $p < .01$, anxious, $r = -.43$, $p < .01$, tense, $r = -.31$, $p < .05$, and guarded, $r = -.40$, $p < .01$, and were less liked, $r = -.37$, $p < .05$, than patients with lower BDI scores.

Because these results suggest that depressive symptomatology plays a role in how PD patients are perceived, we decided to examine whether PD patients who have the same number of depressive symptoms as controls would be perceived negatively. To address this question, controls were compared to 18 PD patients whose BDI score did not exceed the highest BDI score in the control group, six, on three dependent variables: Pers+, Beh+, and Overall+. The means of both groups on all types of ratings are shown in Table 14. One-way univariate ANOVAs[37] with PD status as a between-subjects factor revealed that the Overall+, Pers+, and Beh+ mean ratings of PD participants were significantly lower than those of controls, $F(1, 402) = 36.00$, $p < .01$, $\eta^2 = .08$ for Overall+, $F(1, 402) = 11.88$, $p < .01$, $\eta^2 = .03$ for Pers+, $F(1, 402) = 48.07$, $p < .01$, $\eta^2 = .11$ for Beh+. These results show that the personality and behavior of PD participants with few depressive symptoms were perceived more negatively than the personality and behavior of controls.

Patient characteristics and the effect of LSVT. The magnitude of the LSVT's effect on the way a patient is perceived may depend on patient characteristics. For example, it is

---

37 ANOVAs were chosen for these group comparisons because the tests of skewness and kurtosis showed that the Pers+, Beh+, and Overall+ score distributions of both groups met the normality assumption. Also, modified O'Brien's procedures demonstrated that the homogeneity of variance assumption was satisfied in each group comparison.

possible that the LSVT improves others' perception of male PD patients to a greater extent than that of female patients. Although the main analyses of Hypothesis 2b provided little evidence for such LSVT-related changes in observers' judgments, the exploratory analyses suggested that the ratings of the LSVT, but not ARTIC, group, increased significantly from Time 1 to Time 2. Thus, relationships between the effect of LSVT on observers' ratings and patient characteristics were explored.

The same characteristics as in the preceding subsection (gender, age, handedness, Hoehn-Yahr stage, time since PD diagnosis, education, raw BDI score at Time 1, dichotomized BDI score at Time 1, MMSE score, and number of PD medications taken by the patient) were examined. For the discrete variables of gender, handedness, and dichotomized BDI score, a separate Student's t test with a characteristic as the independent variable and the difference between Time 2 versus Time 1 ratings (i.e., Time 2 - Time 1) as the dependent variable was performed for each rating type (Pers+, Beh+, and Overall+). For all other characteristics, a Pearson correlation was computed between a characteristic and each type of rating. Age and a change in the Beh+ ratings were significantly correlated, $r = .58$, $p = .02$. Also, the change (from Time 1 to Time 2) in the Overall+ ratings of LSVT patients with low BDI scores ($M = 0.49$, $SD = 0.69$) was greater than that in the Overall+ ratings of LSVT patients with moderate BDI scores ($M = -0.16$, $SD = 0.42$), at a marginally significant level, $F(1, 14) = 3.78$, $p = .07$. Similarly, the change in the Pers+ ratings of LSVT patients with low BDI scores ($M = 0.45$, $SD = 0.69$) was greater than that in the Pers+ ratings of LSVT patients with moderate BDI scores ($M = -0.29$, $SD = 0.56$), at a marginally significant level, $F(1, 14) = 4.46$, $p = .053$.

Hypothesis 2c: LSVT-related Improvements in the Negative Perceptions of PD Patients are Mediated by Changes in Facial Mobility

As discussed in the Method section, a six-step statistical procedure (Shrout & Bolger, 2002) was proposed as a powerful test of a mediation model, such as the one predicted by Hypothesis 2c (see Figure 2). In Step 1 of the procedure (i.e., a step in which the association between treatment status and the perception of the participant is explored), a WJ procedure was carried out to contrast the mean change in the ratings of the LSVT group (from Time 1 to Time 2) with that in the ratings of the Untreated group, separately for each rating type (Overall+, Pers+, and Beh+). These contrasts showed that the average change in Pers+ was significantly greater in the LSVT ($M = 0.20$, $SD = 0.80$) than in the Untreated group ($M = -0.03$, $SD = 0.68$), $F(1, 427.54) = 3.90$, $p = .049$, $ES = 0.07$. No additional effects were significant (all p-values > .10).

In Step 2, to determine which a paths (paths indicating relationships between treatments and the mediators [i.e., facial movement measures] in Figure 2) were significant, a WJ procedure was performed to contrast the mean value on a facial movement measure of the LSVT group with that of the Untreated groups, separately for each mediator (AU Lability Δ, AU Variability Δ, AU Complexity Δ). Step 2 revealed that the AU Variability Δ mean was significantly greater in the LSVT ($M = 1.81$, $SD = 3.04$) than in the Untreated group ($M = -0.94$, $SD = 2.33$), $F(1, 28.13) = 8.46$, $p < .01$, $ES = 0.88$. Also, the AU Lability Δ mean was significantly greater in the LSVT ($M = 5.00$, $SD = 8.23$) than in the Untreated group ($M = -1.71$, $SD = 6.82$), $F(1, 30.95) = 7.34$, $p = .01$, $ES = 1.05$. No additional effects were significant (all p-values > .10). Because AU Complexity Δ was not significantly associated with treatment status, it was dropped from the model.

In Step 3, to determine which b paths (paths indicating relationships between the mediators and changes in observers' ratings in Figure 2) were significant, the Overall+ ratings

were regressed onto AU Variability Δ, and AU Lability Δ. The same procedures were separately carried out with the Pers+ and Beh+ ratings. Given the violations of normality assumption in the distributions of the mediator variables (see footnotes 22 and 24), a nonparametric least trimmed squares regression was used for each rating type. As none of the regression analyses were statistically significant (all p-values > .10), the test of Hypothesis 2c was aborted after Step 3, as recommended by Shrout and Bolger.[38] Thus, contrary to this hypothesis, these facial movement measures did not appear to mediate the effect of LSVT on personality and behavior judgments. Although changes on AU Variability and AU Lability did not mediate the effects of LSVT, it is possible that these aspects of facial mobility served as the basis for personality and behavior inferences, but that the magnitude of changes on these variables or the size of the LSVT group were too small to detect a statistically significant effect. To rule out this possibility, the correlations between participants' Time 1 scores on these measures and observers' Time 1 ratings were examined, separately for each type of rating (Pers+, Beh+, or Overall+). None of the correlations were significant (all p-values > .10).

38 As an additional analysis, three mediators in the model were replaced with only mediator, CV_H1, and the six-step procedure testing the model was repeated. The same negative findings in Step 3 were obtained, and the test was aborted.

Discussion

We now turn to a discussion of the study findings. First we will examine the evidence for adequate intercoder reliability. Next, we review findings related to parkinsonian facial movement deficits. Subsequently, we discuss the findings regarding Hypotheses 2a (i.e., that PD participants will be perceived more negatively than controls), 1 (i.e., that LSVT improves facial movement deficits in PD), and 2b (i.e., that LSVT leads to a more positive perception of PD participants by observers), respectively. We then consider why, contrary to Hypothesis 2c, the facial movement measures used in this study did not appear to mediate the effect of LSVT on the perception of LSVT patients. Finally, we suggest possible directions for future research and acknowledge several study limitations.

Indices of Intercoder Reliability

The overall intercoder reliability of .71 in the present study is within the spectrum of reliabilities considered to be acceptable by the creators of FACS (≥.7; Ekman, Friesen, & Hager, 2002a) and is comparable to the reliabilities obtained in other investigations using FACS (e.g., Sayette et al., 2005; Simons et al., 2004; Symons et al., 2010). Although intercoder reliabilities in FACS studies typically exceed .71, it should be kept in mind that only specific AUs were scored in the vast majority of these investigations (for a selection of such studies, see Ekman & Rosenberg, 2005). Given that the coders in the present study scored nearly all AUs, their task was especially challenging. Moreover, because participants' speech production further complicated the coders' scoring, the somewhat low intercoder reliability is not surprising. More importantly, intercoder agreement in AU Lability and AU Variability scores (i.e., .80 and .75, respectively), two dependent measures that yielded significant findings, were well within the acceptable range. This strongly suggests that the significant findings of this study in regards to

facial movement (results showing the presence of facial movement deficits in our sample of PD patients and results supporting Hypothesis 1) were based on valid FACS coding. The fact that, on average, 81% of events scored by the primary coder in a given segment were also scored by the reliability coder also confirms that the total number of events (i.e., AU Lability) was measured reliably in the present study. In contrast, because intercoder agreement in measuring AU Complexity, .60, was unacceptable, the negative findings on this measure can be due to the errors of coding and will not be interpreted in the following discussion.

Besides reliably measuring AU Lability and AU Variability, coders judged facial events to begin and end at roughly the same time, as suggested by the fact that the average difference between the event onset times and event offset times were only 0.17 s and 0.58 s, respectively. Taken together, the measures of intercoder agreement that the significant findings of this investigation, discussed next, are not due to coding errors.

Deficits in Spontaneous Facial Movement of PD participants

We confirmed the presence of deficits in spontaneous facial movement produced by the our sample of PD patients. Specifically, the present study demonstrated that during the production of happy monologues, spontaneous[39] facial activity of PD patients, compared to

---

39 Though participants knew that they were being filmed while producing the monologues, it is improbable that their facial expressions were posed as a result of this knowledge. This is because none of participants were told that LSVT was hypothesized to improve not only their vocal parameters, but also their facial movement. Thus, though demand characteristics (e.g., visible camera, experimenter being present in the room) may have affected participants' behavior overall, it is unlikely that such characteristics strongly influenced their facial movements specifically.

demographically matched controls, showed a decreased repertoire (AU Variability) and rate of change (AU Lability) of facial expressions. These results extend previous investigations showing reduced facial expressivity in PD (e.g., Buck & Duffy, 1980). Although previous findings have demonstrated that PD patients have deficits not only in spontaneous (Simons et al., 2004) but also in posed emotional expression (Borod et al., 1990; Bowers et al., 2006), at least one study (Smith et al., 1996) suggested that deficits in spontaneous expression were especially prominent in PD. This finding confirms the impression that frequently emerges during a clinical observation of a PD patient (e.g., Monrad-Krohn, 1924). The greater impairment of spontaneous facial expression in PD is also consistent with the evidence that neurological mechanisms underlying posed versus spontaneous facial expression are largely distinct (Rinn, 1984). In keeping with this distinction, we next consider this study in the context of other investigations of spontaneous facial expression in PD.

Agreements and discrepancies with previous findings. PD participants in this study were impaired on two measures of facial mobility, AU Lability and AU Variability. To our knowledge, only one previous study (Smith et al., 1996) used a measure comparable to AU Lability -- number of changes of facial expression -- to examine facial expressivity in PD. PD participants in that study produced fewer changes of facial expression than controls. Given a number of methodological differences between the two studies (e.g., in contrast to producing emotional monologues in this study, PD participants in the Smith et al. study viewed video clips), the replication of this finding provides strong empirical evidence for the decrease in number of spontaneous facial expressions in PD relative to controls.

Besides measuring AU Lability, the number of unique facial events[40] produced by participants, AU Variability, was used as a dependent measure in the present study. To our knowledge, no previously published investigations quantified this or similar aspects of facial movement in PD. Thus, our finding of decreased AU Variability levels in PD patients relative to controls extends previous research by suggesting that PD reduces one's repertoire of distinct facial expressions.

Lack of group differences on AU Time. In contrast to the group differences on AU Lability and AU Variability, both PD patients and controls did not differ on "AU Time," a variable which, contrary to our initial plans, was not included in the main analyses of the study. AU Time measured the proportion of time during which one or more AUs -- other than AU 0 which denotes the absence of an observable facial action -- is produced. For instance, if a given participant produced at least one AU for 30 s of a 60-second video clip, the AU Time value for that clip was 0.5. To our knowledge, this or a similar measure was not included in any published studies of facial movement in PD.

Because the preliminary examination of the video clips revealed that AU Time had little relation to overall facial mobility, it was not included in the analyses reported in the Results section. Specifically, some participants in each group produced the same AU(s) for the duration of the entire monologue. As a result, despite the high AU Time value of these participants, their face had a mask-like, permanently "frozen" appearance. Conversely, some participants who produced frequent but short facial movements (e.g., quick brow raises [AU 1+2] or eyelid tightening [AU 7]) had moderate AU Time values and faces that appeared animated. Nevertheless, we decided to examine whether the PD participants and controls differed on AU

---

40 A facial event is a set of one or more AUs elicited by the same stimulus.

Time via a post-hoc Mann-Whitney U test. Confirming our intuitions, the average AU Time levels for PD patients (M = 0.60, SD = 0.42) and controls (M = 0.63, SD = 0.39) did not differ significantly, $p = .748$. Given that both group means for AU Time were not near either 0 (the lowest possible AU Time value) or 1 (the greatest possible AU Time value), it is unlikely that this negative finding is a result of floor or ceiling effects. Rather, it appears that the changes in facial expressions, which were measured by AU Lability and AU Variability, rather than their presence, which was measured by AU Time, distinguish PD patients from demographically matched controls.

To summarize, the present study demonstrated that during the production of happy monologues, PD patients, compared to demographically matched controls, showed a decreased repertoire (AU Variability) and rate of change (AU Lability) of spontaneous facial expressions. In contrast, the proportion of time during which a facial expression was observed was similar in both groups. These findings replicate previous findings and add to our knowledge about parkinsonian deficits in spontaneous facial movement.

Hypothesis 2a: Negative Bias in Others' Perception of PD patients

In support of Hypothesis 2a, observers judged both the personality and behavior of PD participants more negatively than those of demographically matched controls. Specifically, PD participants compared to controls were judged to have more introverted and anxious personalities. Further, relative to controls, PD patients appeared bored, tense, sad, and guarded and were less liked. The subclinical depression of some PD participants (i.e., because patients with BDI scores indicative of severe depression were excluded from this study, the depressive symptoms of some PD patients were likely insufficient for a clinical diagnosis of depression) does not explain these negative impressions, as PD participants with minimal depressive

symptoms were also perceived negatively. Nonetheless, subclinical perception in PD appeared to contribute to the negative perception of PD patients, as suggested by the significant negative correlations between PD participants' scores on the Beck Depression Inventory (BDI) and observers' personality (Pers+) and behavior (Beh+) judgments of these participants. These correlations reflected observers' tendency to perceive PD participants with higher BDI scores as more introverted, dependent, anxious, tense, and guarded, and to like them less than PD participants with lower BDI scores.

The perceived personality and behavior characteristics of PD patients are the same or closely related to the ones on which PD participants, relative to controls, were judged negatively in two previous investigations (Pentland et al., 1987, 1988). In these studies, PD participants were judged to be less likeable and more anxious, tense, hostile, suspicious, and unhappy than controls. Unfortunately, because no personality measures were administered, the current study cannot investigate the accuracy of observers' negative impressions. It is possible that these impressions reflected genuine personality differences between PD participants and controls. To clarify this issue, evidence from other studies of how PD patients are perceived is considered next.

Evidence for a negative bias in others' perception of PD patients. There is a substantial amount of evidence that PD patients tend to exhibit characteristic personality traits. Almost a century ago, Camp (1913) noted that PD "affected mostly those persons whose lives had been devoted to hard work, either mental or physical," and that individuals with this disease had "a very small proportion of alcohol or tobacco users." Subsequent case reports have elaborated on this idea noting that people with PD were frequently law-abiding and diligent individuals with a perfectionistic, overly moralistic, and rigid approach to life, who showed an

excessive amount of self-control and little spontaneity of thought (e.g., Korten & Ketterings, 1972, as cited in Todes & Lees, 1985; Sands, 1942). In addition, these individuals tended to cope poorly with emotional stress and had a manic-depressive quality to their mood patterns (Prick, 1966, as cited in Todes & Lees, 1985). Based on the review of such reports, Todes and Lees (1985) concluded that the personality and behavior of PD patients was frequently characterized by restraint, anhedonia, emotional inflexibility, introspection, and suppressed aggression. However, because the evidence reviewed by Todes and Lees primarily included clinical observations of very few patients, subsequent studies used more methodologically rigorous methods to ensure that the authors' conclusions were generalizable and largely independent of clinicians' biases.

These more recent investigations have demonstrated that PD patients score lower on the novelty seeking scale and higher on the harm avoidance scale of the Tridimensional Personality Questionnaire than do healthy controls (Fujii, Harada, Ohkoshi, Hayashi, & Yoshizawa, 2000; Kaasinen et al., 2001) or patients with rheumatism (Menza, Golbe, Cody, & Forman, 1993). Given that individuals scoring low on the novelty seeking scale of that personality measure tend to be reflective, rigidly moral, stoic, nonimpulsive, frugal, orderly, and persistent (Menza et al., 1993), these findings are in line with the earlier observations of the parkinsonian personality and behavior. Notably, because increased harm avoidance is associated with social shyness, excessive worrying about the future, and depression, the higher mean score of PD patients on that scale is consistent with Prick's (1966, as cited in Todes & Lees, 1985) report on the emotional difficulties these individuals experience. In addition, PD patients were shown to be more nervous, quiet, serious, introverted, and self-controlled, on average, than their monozygotic twins without PD (Duvoisin, Eldridge, Williams, Nutt, & Calne, 1981; Ward et al., 1984).

Intriguingly, a prospective study (Bower at al., 2010) of a large cohort of 6822 participants showed that high scores on the anxious and pessimistic personality subscales of the Minnesota Multiphasic Personality Inventory (MMPI) were associated with increased risk of PD even when the MMPI was completed at ages 20 to 39 (i.e., more than two decades earlier than the mean age of PD onset), strongly suggesting that high scores on these subscales were a premorbid aspect of the disease. In contrast, scores on the introversion, depressive, and novelty seeking subscales of the MMPI did not predict risk for PD, suggesting that introversion and depression may emerge as a result of disease progression (Arabia et al., 2010; Bower at al., 2010).

Thus, there is considerable evidence that certain personality and behavior characteristics, including moral rigidity, perfectionism, cautiousness, nervousness, excessive self-control, excessive worrying, depressive symptoms, and introversion, tend to occur more frequently among individuals with PD than individuals without PD. Moreover, as many of these characteristics (e.g., nervousness, introversion, and excessive worrying) are closely related to the dimensions on which observers in this study made personality and behavior judgments (e.g., sad versus happy, introverted versus extraverted, tense versus relaxed), this evidence suggests that observers' negative impressions are at least partly accurate. However, the possibility of some bias in observers' judgments remains. For example, it is possible that a slightly introverted and anxious patient seemed to be very introverted and extremely tense to observers. As a result, observers' judgments may have both reflected the patient's personality and behavior and showed a negative bias. To determine whether the perception of individuals with PD is biased, several studies examined whether PD patients and non-PD controls who scored similarly on personality measures were nevertheless perceived differently by observers.

For example, Pentland and colleagues (1987) were the first to compare PD patients and controls (demographically matched patients with cardiac ischemia) -- two participant groups who scored similarly on personality measures -- in terms of how they were perceived by observers. Observers, blind to whether a given patient had PD, rated each patient on a number of personality and behavior scales based on a silent video clip from a clinical interview. PD patients were rated more anxious, tense, hostile, suspicious, and unhappy than cardiac patients. Pentland and colleagues (1988) then showed that despite their experience with PD patients, healthcare professionals also perceived these individuals with a negative bias. As suggested by a recent study (Hemmesch, Tickle-Degnen, & Zebrowitz, 2009), this bias distorts the perception of not only healthcare professionals, but also age peers of PD patients. After viewing the video clips of PD patients describing a positive experience, age peers judged patients with a high degree of masking to be less socially supportive than those with a low degree, despite the fact that both groups scored similarly on a measure of social supportiveness. This finding also suggests that the negative bias is associated with a parkinsonian problem in nonverbal communication, facial masking. Given the evidence that facial masking diminishes when PD patients speak about positive topics (Takahashi, Tickle-Degnen, Costa, & Latham, 2010), the persistence of this bias in the Hemmesch et al. study underlines its tenacity.

Taken together, all of the reviewed evidence suggests that observers' negative impressions of PD patients in the present study reflect the effect of two different factors: genuine personality characteristics more commonly found in this clinical population and a negative bias of observers. Because participants' personalities were not formally assessed, the relative magnitude of each effect is unknown. Based on the available evidence, our preliminary conclusion is that the lower Pers+, Beh+, and Overall+ ratings of PD patients, compared to

controls, partially reflect the negative bias in the perception of individuals with PD. Importantly, as shown by observers' negative perception of PD patients with minimal depressive symptoms (i.e., patients with BDI scores of less than seven), the subclinical depression of some PD patients does not fully explain these negative impressions. However, as we discuss next, PD patients' depressive symptoms contribute to the negative way in which they are perceived.

Perception of interindividual differences in PD patients' personality and behavior. Because personality assessment was not carried out in this study, any conclusions regarding the accuracy of observers' personality judgments must remain speculative. However, the negative and significant correlations of observers' Pers+ and Beh+ ratings with participants' BDI scores provide preliminary evidence that observers were sensitive to interindividual variation in PD patients' mood. Moreover, observers' tendency to judge PD patients with higher BDI scores as anxious, tense, guarded, and introverted may mean that despite perceiving PD patients with a negative bias (see the preceding subsection), observers are nevertheless sensitive to the symptoms of depression occurring in this clinical population. Given the estimate that 67% of depressed PD patients have a comorbid anxiety disorder (Menza, Robertson-Hoffman, & Bonapace, 1993), it is likely that many PD patients with higher BDI scores tended to be more anxious than those with lower scores. Also, in light of the evidence that hypervigilance is a core component of anxiety disorders (Beck, Emery, & Greenberg, 2005), it is conceivable that PD patients with high scores on the BDI tended to be more tense and guarded than those with low scores. Further, as low scores on extraversion have been linked to depressive symptomatology and diagnoses (for reviews, see Clark & Watson, 1991a&b), PD patients with higher BDI scores

probably tended to be less extraverted than those with lower scores. In sum, these studies suggest that observers in this study were sensitive to participants' depressive symptoms.[41]

Given the significant correlations of BDI scores with Pers+ and Beh+ ratings, it is worth considering why similar relationships between patients' scores on the measures of personality or mood and observers' judgments were not obtained in previous studies (Lyons et al., 2004b; Tickle-Degnen & Lyons, 2004). These studies showed that observers' ratings had moderate sensitivity to PD participants' levels of conscientiousness, agreeableness, and openness to experience and low sensitivity to participants' levels of neuroticism and extraversion. Moreover, the judgments of neuroticism and extraversion were based on facial masking and blinking rate -- two aspects of facial movement that are heavily affected by PD (Deuschl & Goddemeier, 1998; Karson, 1983) and are not associated with one's neuroticism or extraversion levels (Lyons et al.,

[41] Alternatively, it may be that observers' ratings were not based on the signs of subclinical depression, but on some cues of behavior or appearance that are usually unrelated to depression. These cues may have caused PD participants who were rated negatively in this study to appear in a negative light outside the laboratory as well. These negative perceptions may have amounted to stigmatization, whether actual or perceived by the patient, which contributed to patients' depressive symptomatology. Findings showing that stigmatization levels and social isolation account for a large proportion of variance in the BDI scores of PD patients (Karlsen, Larsen, Tandberg, & Mæland, 1999; Schrag, Jahanshahi, & Quinn, 2001; for review, see Frisina, Borod, Foldi, & Tenenbaum, 2008) support this alternative hypothesis. Although this account is less parsimonious than the one suggesting a direct link between patients' depressive symptoms and observers' ratings, it cannot be ruled out.

2004b). Because neuroticism levels are positively associated with depressive symptoms (Weiss et al., 2009), the present correlations between BDI scores and observers' ratings conflict with these previous findings which show observers' apparent lack of sensitivity to interindividual variation in neuroticism levels among PD participants.

One reason for the discrepant findings may be the fact that PD participants in these previous studies, as compared to the ones in this study, had much lower levels of depressive symptomatology. Specifically, a cutoff of 6/7 on the Short Form of the Geriatric Depression Scale -- an exclusion criterion that would screen out some PD patients with mild levels of depressive symptoms (Almeida & Almeida, 1999) -- was used by Tickle-Degnen and colleagues. In contrast, four PD patients in our study had BDI scores above 19 suggestive of moderate levels of depressive symptoms (Beck, Steer, & Brown, 1996). As a result, the outward manifestations of subclinical depression among PD patients of this study may have provided an additional information source for observers and thus may have improved the sensitivity of observers to interindividual differences. Conversely, the absence of such cues in nondepressed PD patients in previous studies may have lowered observers' sensitivity. Thus, based on the preliminary results of the present study, we hypothesize that others' perception of PD patients may be more sensitive to interindividual variation than previously thought as long as a naturalistic sample of PD patients (i.e., one that includes PD patients with comorbid depressive and anxiety symptoms) is examined.

To summarize, results suggest that PD patients with subclinical depression are perceived more negatively than nondepressed PD patients. However, in contrast to PD which leads to perceptual biases (see previous subsection), subclinical depression may help observers detect interindividual differences in the personality and behavior of PD patients. Because no

personality measures were administered in this study, and other studies did not examine the effect of clinical or subclinical depression on the perception of PD patients, no direct evidence for this hypothesis is currently available. Clearly, the present findings underscore the need for future investigations clarifying the role of subclinical depression, a condition that is highly comorbid with PD (Burn, 2002; Leentjens, 2004), in the social functioning of PD patients.

One may expect that similar to depressive symptomatology, the race of observers and participants affects the judgments of participants' personality and behavior. Because only six participants out of 56 participants in this study were non-Caucasian (see Table 3), this study did not examine the effect of participant race on such judgments. In contrast, the multi-racial composition of the observer sample (see Table 4) enabled us to show that the race of observers did not moderate their perception of participants. These findings are consistent with the empirical evidence for agreement between the judgments of individuals' personality made by observers of their own race and culture and those made by observers of a different race and culture (Albright et al., 1997; Rule et al., 2010), regardless of whether the judged individuals were Chinese, American, or Japanese. Thus, we hypothesize that the present findings can be replicated with observers and participants of different races and ethnicities.

Psychosocial impact of PD patients' problems in verbal and nonverbal communication. We have seen that nonverbal communication problems in PD, such as facial masking, can lead to negative bias and lack of sensitivity in observers' perceptions, which suggests that these problems may also disrupt the interactions of PD patients outside the laboratory. However, with the exception of the studies reviewed earlier (e.g., Lyons et al., 2004b; Pentland et al., 1987), these problems have been primarily discussed from a clinician's perspective. Specifically, several authors (Leentjens, 2004; Marsh, 2000) have commented that

due to being a symptom in both PD and depression, facial masking, along with other symptoms shared by these illnesses, complicates the diagnosis of depression in PD, potentially leading to both misses and false alarms (Lieberman, 2006). Studies (Shulman, Taback, Rabinstein, & Weiner, 2002; Weintraub, Moberg, Duda, Katz, & Stern, 2003) showing poor accuracy of around 35% in the diagnosis of depression among PD patients strongly support this conjecture. Given that early detection of depression may improve the cognitive performance of PD patients (Kuzis, Sabe, Tiberti, Leiguarda, & Starkstein, 1997), the deleterious impact of facial masking and other overlapping symptoms of PD and depression cannot be overstated.

In addition to making the diagnosis of depression in PD problematic, nonverbal communication deficits may have other negative effects on the lives of PD patients (e.g., stigmatization). Although none of the known studies distinguished between verbal and nonverbal communication when examining this issue, a clear distinction may often be impossible, as the production and comprehension of two communication types are usually interdependent (Rosenblum, Miller, & Sanchez, 2007). Thus, despite not being specific to nonverbal communication problems in PD, these study findings may provide clues about the impact of facial masking on the lives of people with PD and their caregivers.

When asked about PD-related problems in their lives, PD patients have indicated a range of problems in both surveys (Ellgring et al., 1993; Jenkinson, Peto, Fitzpatrick, Greenhall, & Hyman, 1995) and intensive interviews (Dakof & Mendelsohn, 1989; Marr, 1991). Extending well beyond the motor and cognitive impairments, these problems included limitations in social interaction, communication difficulties, loss of social support, and social stigma. A factor analysis on PD patients' responses to a survey showed that these problems could be classified into three categories: physical, cognitive, and social (Brod et al., 1998). Communication

difficulties (e.g., “expressing feelings”) fell into the social group, which also included “feelings of social isolation,” “feelings of being a burden,” “feelings of depression,” and “being taken seriously.” Despite not allowing any causal interpretations, this finding suggested that communication difficulties were closely related to the disruptions in the social functioning of individuals with PD. Moreover, such disruptions were widespread, as even the least common one (“being taken seriously by others”) was endorsed by 24% of the respondents.

Although not as readily quantifiable as survey data, PD patients’ responses in semi-structured interviews often provide a more nuanced view of the impact of communication difficulties in PD. A study (Miller et al., 2006) using this method to explore the psychosocial consequences of speech changes in PD showed that PD patients were less concerned with disturbances in the acoustic parameters of their speech than with the effect these disturbances had on their interaction with others. One prominent concern that appears equally relevant to both verbal and facial communication was the frustration of not being able to express one’s needs and, as a result, being frequently ignored or stigmatized (e.g., “considered stupid”) by others. Patients’ ways to manage the problems they perceived were largely passive including taking on the role of a listener or even complete withdrawal from a social interaction. Fortunately, active ways of coping (e.g., explaining to people their communication difficulties, keeping their sentences short, or seeking alternative communication channels, such as email or note writing) occurred as well. Given that exclusion from conversations undermines the dignity of the elderly (Woolhead, Calnan, Dieppe, & Tadd, 2004), these findings underscore the major impact of communication problems on the psychological well-being of PD patients.

Empirical studies (Miller, Noble, Jones, Allcock, & Burn, 2008; Schrag et al., 2001) have confirmed the conclusion based on the interview data that it is often a patient’s perception, rather

than a clinician's assessment, of these communication problems that may be most important when considering the disruption to a patient's life. Specifically, these studies revealed moderate and significant correlations (in the .30 - .39 range) between the self-reported extent of patients' communication problems and self-reported disease severity and levels of depression. Yet, it is the subjective dimension of a patient's communication problem that is most difficult to capture within the time constraints of a medical appointment. This is demonstrated by the finding that the extent of verbal communication problems, as assessed by patient self-report (Miller et al., 2008) or listeners' intelligibility ratings (Miller et al., 2007), is only weakly correlated with clinician-rated disease severity measures (e.g., Hoehn and Yahr stage) or the diagnostic assessments of voice and speech. In contrast, PD patients and their caregivers tend to have similar perceptions of the severity of a patient's communication problems (Miller et al., 2008). Altogether, these findings suggest that these perceptions, although possibly inaccurate, need to be taken into account during clinical decision-making because they determine how PD patients and caregivers respond to their communication difficulties.

The important role of PD patients' subjective views of their communication difficulties, suggested by these results, raises the question of whether impairments in nonverbal communication, the subject of interest in this study, are as evident to the patient, as the impairments in voice and speech, the main focus of some studies (e.g., Miller et al., 2007) just discussed. Because one usually does not see his or her own facial movement, it may be argued that PD patients are relatively unaware of their facial masking. This lack of awareness may thus reduce patients' embarrassment about this problem -- a sentiment they share in regards to their speech (Miller et al., 2006) -- and consequently reduce its psychosocial aspect.

Two pieces of evidence challenge this argument. First, individuals with PD may be aware of their facial expressivity deficits. Although the self-reported facial expressivity levels of PD participants were similar to those of control participants in one study (Simons et al., 2004), PD participants' self-reported facial expressivity was significantly lower than that of controls in another study (Mikos et al., 2009). In addition, the proportion of participants substantially over- or under-reporting their facial expressivity was similar for both groups in the Mikos et al. study. Second, even if individuals with PD are relatively unaware of their facial movement deficits, these problems may nonetheless have an indirect effect on patients' well-being via straining their interactions with close others. To test this hypothesis, future research needs to extend the studies of Pentland, Tickle-Degnen and colleagues (Hemmesch et al., 2009; Pentland et al., 1998), which have shown that facial masking leads to perceptual biases in the laboratory, to more naturalistic settings.

To summarize, despite the lack of evidence directly linking nonverbal communication deficits with PD patients' quality of life, extant studies suggest that these impairments may affect the psychosocial functioning of this clinical population in several ways. Specifically, as a result of causing a misdiagnosis of depression, facial movement deficits may delay the treatment of depression in PD patients worsening their cognitive decline. Furthermore, facial masking leads to a negative bias in the perception of PD patients by others. It remains to be seen whether this bias persists across various interactions outside the laboratory, amounting to the stigmatization of a patient. Finally, PD patients' subjective perception of their communication deficits is associated with the severity of depression reported by these individuals. Given that many, if not all, PD patients are aware of their facial expressivity deficits, this association probably holds true for impairments in both verbal and nonverbal communication.

Hypothesis 1: Effects of LSVT on PD Patients' Facial Movement

In support of Hypothesis 1, study results showed that LSVT, relative to control articulation therapy, can improve facial movement in PD. Specifically, the improvement of LSVT patients on the canonical variate combining facial variability and lability (i.e., CV_H1) was significantly greater than that of the control therapy group (ARTIC) or any other groups. Although the direct comparisons of LSVT and ARTIC groups on AU Variability Δ and AU Lability Δ did not reveal statistically significant differences, comparisons of both groups with the Untreated patients bore out the superiority of LSVT over the ARTIC therapy in improving facial movement. Specifically, whereas the increase of LSVT patients on either variable was significantly greater than the corresponding increase of Untreated patients, the differences between ARTIC and Untreated patients on these variables were not significant. Similar results were obtained when the relative measures of changes in facial mobility were used (i.e., with the change of each participant measured as the proportion of the average facial mobility of that participant across Times 1 and 2). These differences between ARTIC and LSVT groups showed that LSVT patients' facial mobility increased due to treatment and not as a consequence of a more general effect, such as improved mood as a result of gaining a therapist's attention.

Given the encouraging evidence that LSVT reduces facial movement deficits in PD, it is important to examine whether certain PD patients are better suited to benefit from this treatment than others. The present study revealed one finding that may be relevant to this issue: an interaction between the size of LSVT-related effects on facial movement and a participant's education level. Specifically, LSVT patients with a higher education level demonstrated a greater increase in facial movement, as measured by CV_H1, than participants with a low education level. Although one speech language pathologist noted that she adapted the

application of LSVT according to patients' education level (Parkinson's Companion, 2010), this finding, to our knowledge, has not been noted in prior studies involving the effects of LSVT. This is a surprising result because LSVT is administered primarily by asking patients to imitate a therapist's voice and speech, and involves little explicit instruction (Trail et al., 2005). Because the cognitive load needed for such imitation is relatively low, it is not clear why the patient's education level played a role in the magnitude of LSVT's effect.

However, it may be that even the relatively low cognitive demands of LSVT exercises are processed differently depending on the educational attainment of the patient. The negative association between educational attainment and the risk of dementia in PD (Glatt et al., 1996) supports this interpretation. This interpretation is also consistent with the fact that education level is one of the strongest predictors of cognitive reserve - the ability to optimize or maximize performance through differential recruitment of brain networks (Stern, 2002). Thus, due to higher cognitive reserve, PD patients with higher education levels in our sample may have been able to process LSVT assignments more efficiently and, as a result, benefit from LSVT more than PD patients with lower education levels. Because the relationship between the effect of LSVT on facial movement and a participant's education level was obtained in exploratory analyses without corrections for multiple comparisons, this intepretation is speculative and needs to be tested in future studies.

Speculative mechanisms underlying the effect of LSVT on facial movement. What is the mechanism underlying the effect of LSVT on facial movement? Although direct evidence regarding this issue is lacking, hypotheses can be formulated based on the emerging knowledge about neuroplasticity in the parkinsonian brain. Investigations in this area not only reveal that non-surgical interventions can slow the progress of PD but also provide information about the

critical components of such interventions. Based on these studies, it is argued that three aspects of LSVT are critical for alleviating facial movement deficits in PD. We now review these components and the evidence supporting their therapeutic potential.

LSVT is intensive and directly targets the deficits of PD patients. PD disrupts the neural programming of motor and vocal output. LSVT is based on the hypothesis that this impairment results in vocal characteristics (e.g., loudness) that are too low to achieve the speaker's objectives (Trail et al., 2005). In view of this hypothesis, it may appear that successful treatment of PD patients' voice and speech deficits should not be based on practicing adequate loudness, but rather on providing compensatory strategies. However, recent studies showed that the neuroplasticity of basal ganglia pathways and therefore the potential to reduce neurological deficits in PD may be higher than previously thought. These reports suggest that the key to unlocking this potential may be the intensive exercise of a compromised brain function.

The initial evidence for this idea came from investigations examining the effects of intensive physical exercise on rodents that were injected with 6-hydroxydopamine (6-OHDA) and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), dopamine neurotoxins, in the medial forebrain bundle (Tillerson et al., 2001; Tillerson, Caudle, Reveron, & Miller, 2002, 2003). Rodents injected with these neurotoxins, which drastically deplete nigrostriatal dopamine levels and cause akinesia of the forelimbs, serve as the animal model of PD. In a pioneering study (Tillerson et al., 2001), rats who were forced to use the impaired forelimb for a week within 12 hours following a 6-OHDA injection (i.e., by putting their nonimpaired forelimb in a cast) did not show preferential use of any forelimb after removal of the cast. Moreover, the striatal dopamine levels of these rats were not significantly different from those of control rats that received sham injection. In contrast, rats who did not wear a cast preferred using an impaired

limb for the entire study duration. Also, akinesia and striatal depletion were evident in the "noncasted" rats 67 days after the injection. Interestingly, the dopamine cells of rats whose nonimpaired limb was restrained three days after the injection were not spared. Similar results were obtained when mice that were forced to run on a treadmill twice a day for 10 days following a bilateral MPTP injection were compared to control mice that did not have treadmill training (Tillerson et al., 2003). The important clinical implication of these findings is that practicing motor skills mildly impaired by PD may considerably slow further declines of those skills and have neuroprotective effects.

Extending these results, subsequent studies (Fisher et al., 1994; Petzinger et al., 2007) showed that lower running speed, an ordinarily permanent motor deficit of mice injected with MPTP, could not only be prevented but reversed by physically training the mice. If the exercise program was long and intensive (Fisher et al., 1994), the decrement in speed was almost entirely eliminated even if the mice were initiated on an exercise program four days after the injection, when the process of MPTP-induced cell death had ended. In other words, the effects of exercise were not only neuroprotective, but also neurorestorative. The importance of exercise intensity became especially apparent in the emerging studies of the benefits of physical training in PD patients. Similar to the exercise regimens administered to rodents, training in these studies involved practicing the abilities compromised by PD (e.g., walking) rather than teaching compensatory strategies.

In two studies that examined the effects of physical training intensity on therapeutic gains in PD patients (Fisher et al., 2008; Pohl, Rockstroh, Ruckriem, Mrass, & Mehrholz, 2003), all basic parameters of gait and balance improved after both high- and low-intensity treadmill training programs. However, the greatest evidence for improvement was observed after the

high-intensity treadmill programs, as compared to the less intensive programs, in both studies. In addition, in the Fisher et al. study, high-intensity but not low-intensity treadmill training resulted in the decrease of corticomotor excitability -- measured via transcranial magnetic stimulation -- which is abnormally high in PD. Other studies showed that besides increasing walking speed, intensive speed-dependent treadmill training reduced postural instability and fear of falling in PD (Cakit, Saracoglu, Genc, Erdem, & Inan, 2007; Herman, Giladi, Gruendlinger, & Hausdorff, 2007). It has been hypothesized that these therapeutic benefits are mediated by increases in brain-derived neurotrophic factor (BDNF), a molecule implicated in neurogenesis and reduction of free radical formation, an important process in the pathophysiology in PD (Archer, Fredriksson, & Johansson, 2010; Keus, Bloem, Hendriks, Bredero-Cohen, & Munneke, 2007).

Altogether, investigations just reviewed demonstrate that physical exercise can halt and even reverse the loss of motor functions in PD. Importantly, rather than being taught strategies to compensate for their deficits, animal or human participants in each of these investigations were intensively trained to use functions (e.g., walking) that were impaired by the loss of dopamine in the basal ganglia. Similar to these training programs, LSVT provides intensive practice of the ability impaired by PD: phonating and speaking with adequate vocal amplitude. As the studies reviewed in this subsection suggest, this therapeutic approach may partly reverse these impairments. However, given that LSVT is a voice treatment, the reviewed evidence does not explain why in addition to diminishing the deficits of hypokinetic dysarthria (Yorkston et al., 2007), LSVT also reduced facial movement deficits of patients in this study. It is hypothesized that at least two factors facilitate the carryover of LSVT's effects on facial movement. Together with the supporting evidence, these factors are discussed in the next two subsections.

LSVT is parameter-specific rather than task-specific. Many current rehabilitation programs for PD patients focus on helping patients overcome motor difficulties through the use of problem-specific therapeutic approaches. Thus, PD patients may practice whole-body movements to alleviate the freezing phenomenon when they intend to start walking (Yekutiel, Pinhasov, Shahar, & Sroka, 1991) and they may be taught the use of external cues for improving their gait once they walk (Suteerawattananon, Morris, Etnyre, Jankovic, & Protas, 2004). Although such problem-specific approaches are effective (e.g., Mohr et al., 1996), there is little evidence that the resultant therapeutic gains generalize across tasks and contexts (Farley, Fox, Ramig, & McFarland, 2008). Due to this lack of generalization, PD patients need to be taught a large number of therapeutic tools if the full range of their motor problems is to be addressed. However, because, due to the constraints on patients' time and cognitive capacity, this is rarely possible, rehabilitation paradigms based on such approaches are unlikely to have a wide-ranging impact on PD patients' motor function.

Instead of teaching PD patients problem- or task-specific strategies, LSVT aims to target abnormalities in the neural programming of vocal amplitude (loudness). Neurophysiologic studies show that similar abnormalities in the programming of movement amplitude are a critical component of a wide range of motor deficits in PD (Berardelli, Rothwell, Thompson, & Hallett, 2001). Further, the programming of amplitude (also known as "amplitude scaling") appears to be dependent on the same brain region, a functionally segregated circuit in the basal ganglia involving the caudolateral regions of the striatum (Jueptner, Frith, Brooks, Frackowiak, & Passingham, 1997; Lehericy et al., 2005), for a wide range of automatic behaviors. Given that voice modulation and spontaneous facial movements are largely automatic, this evidence strongly suggests that their regulation may rely on the same computational operations in the

basal ganglia. This is also suggested by the hypometria, inadequately low ("underscaled") amplitude, of facial movement in PD (Smith et al., 1996). Thus, by targeting the programming of movement amplitude, LSVT may have partly normalized these operations resulting in the reduction of both voice and facial movement deficits in PD. The increased amplitude after LSVT would result in greater AU intensity (i.e., reduced hypometria of facial movement), which may explain some of the findings in this study, as the following hypothetical example demonstrates.

Suppose that before the onset of PD, a given individual had a tendency to produce two facial events while telling a happy story: a high-intensity AU 6 + 12 and low-intensity AU 17. Further, suppose that due to PD, the amplitude of muscle contractions needed to produce these facial events became underscaled. As a result, the motor commands to produce a low-intensity AU 17 would not result in an observable AU, and AU 6+12 would be produced with low rather than high intensity. If LSVT normalized facial movement by rescaling the underlying motor programs, this person would produce two unique events (AU 6+12 and AU 17) after LSVT instead of just one event (AU 6+12) before it and thus show an increase in AU Variability. Further, due to the repeated occurrences of AU 17, AU Lability would increase as well. The mechanism of LSVT's action hypothesized in this scenario leads to a number of testable predictions (see Future Directions below).

Nevertheless, if the rescaled amplitude of automatic movements following LSVT were the only reason PD participants' facial mobility improved after treatment, all automatic movements of these patients rather than only facial movements would have improved. Because this probably did not occur, there must be factors that facilitated the carryover of LSVT's effect to facial movement but not to other automatic behaviors of LSVT patients. We hypothesize that

one of these factors is the close relationship between vocal parameters and the contractions of the lower face muscles. The evidence for this hypothesis is reviewed next.

The contractions of the lower face muscles are closely coordinated with vocal amplitude, the target of LSVT. Studies discussed in the Introduction have shown various instances of coordination between vocal parameters and lip movements, including strong correlations between the timing of vocal fold closure and timing of lower lip movements (Gracco & Löfqvist, 1994) and between vocal intensity and lip velocities (Dromey & Ramig, 1998). Also, significant correlations between vocal intensity and lip velocities and between fundamental frequency (i.e., an acoustic characteristic that underlies pitch and is determined by laryngeal movements) and lip velocities were obtained in 50-86% of participants with normal speech (McClean & Tasko, 2002). Taken together, these results show that vocal parameters (e.g., fundamental frequency) are highly coordinated with those of lip movements.

Complementing these results, additional investigations revealed that similar to lip movements, cheek movements are also associated with vocal parameters. Thus, Yehia, Rubin, and Vatikiotis-Bateson (1998) found that 73% of the variance in the spectral envelope (a measure derived from sound frequency and amplitude) of a sentence was accounted for by the movement parameters of the lips, chin, and cheeks. In another study, correlations of 0.5 - 0.85 were found between the distance traveled by the cheek markers and the spectral envelope of a vowel (Jiang, Alwan, Bernstein, Keating, & Auer, 2000). The exact magnitude of the correlation depended on a speaker, showing the importance of individual differences in these relationships. To our knowledge, associations between vocal parameters and facial movements above the cheeks have not been investigated. Nevertheless, the positive correlations between the movement parameters of muscles in the lower face and vocal amplitude and/or frequency point

to the close relationship between two channels of expression (facial and vocal) and predict that LSVT-related increases of vocal amplitude will be associated with the amplification of some facial movements. Thus, due to its focus on vocal amplitude, LSVT may be effective at changing the facial movement dynamics in PD.

Therapeutic implications. Although this study provides evidence supporting the effectiveness of LSVT for reducing facial movement deficits in PD, it should be kept in mind that due to being a vocal treatment, LSVT is unlikely to be the optimal treatment for these deficits. Instead, a treatment designed to alleviate masked facies that is based on the principles of LSVT may be the best way to improve the facial movement of PD patients. The preceding discussion suggests these principles: high treatment intensity, direct targeting of parkinsonian deficits, and focus on movement amplitude. The development of LSVT BIG, a therapy for alleviating hypometria in PD, shows that the design of an effective new treatment based on these principles is possible (Farley et al., 2008). Instead of learning compensatory strategies for hypometria, LSVT BIG patients practice performing large-amplitude whole-body movements with high effort and intensity inside the therapist's office and in their daily lives. LSVT BIG patients showed a significantly greater improvement on the Unified Parkinson's Disease Rating Scale and timed walking tasks than control PD patients (Ebersbach et al., 2010). As the results of the present study suggest, the principles of LSVT and LSVT BIG may also be critical for treating facial movement deficits in PD.

Hypothesis 2b: The Effect of LSVT on the Perception of PD Patients by Others

The main analyses of this study provided little evidence for Hypothesis 2b that LSVT, relative to ARTIC therapy, will improve others' perception of PD patients. Specifically, changes in the ratings of LSVT participants at Time 1 versus Time 2 did not differ significantly from

those in the ratings of the ARTIC participants (Contrast 2) or those in the ratings of three other groups combined (Contrast 1). This was the case for the Pers+, Beh+, and Overall+ ratings. However, additional analyses qualified these results showing significant increases in the Overall+ and Pers+ ratings of the LSVT group from Time 1 to Time 2. This preliminary result suggests that LSVT may have improved observers' perception of LSVT patients' personality. Intriguingly, the Overall+ and Pers+ ratings of the Control group showed similar increases. By contrast, changes in the Beh+ ratings of any participant group were not significant.

Potential reasons for the obtained results. Although the main contrasts did not support Hypothesis 2b, the present findings, when considered together, do offer preliminary evidence for this hypothesis. Specifically, the significant increases (from Time 1 to Time 2) in the Pers+ ratings of the Control group clearly contributed to the lack of statistical significance of Contrast 1 (the change in the ratings of LSVT participants versus changes in the ratings of three other groups) in the main analyses. Although Contrast 2 (the ratings of ARTIC versus LSVT patients) was also nonsignificant, additional analyses showed that there was an important difference between the two groups: whereas the improvement in observers' Pers+ judgments of LSVT patients was significantly greater than the corresponding changes in the perception of the Untreated group, changes in the Pers+ ratings of the ARTIC and Untreated patients did not differ significantly. In other words, whereas LSVT, relative to no treatment, apparently led to the improved perception of patients' personalities by observers, this was not the case for the articulation therapy. Furthermore, the analyses of ratings on each personality scale showed that LSVT patients were judged to be more secure, extraverted, and independent, and were more liked after than before treatment. Because all these traits are considered to be positive attributes by most young Americans (Cottrell, Neuberg, & Li, 2007), such as the ones in our observer

sample, these results offer preliminary evidence that LSVT may help PD patients to be perceived in a more positive light.

Because no treatment was administered to controls during the study, the significant increase in the Pers+ ratings of these participants is puzzling. Because each observer never viewed more than one clip of a given participant, this result cannot be due to the exposure effect (i.e., tendency to perceive a stranger in a more positive light upon repeated exposure). Instead, we hypothesize the following explanation for this unexpected finding. It may be that due to their increased familiarity with the study procedures and personnel, all participants felt more comfortable and confident while producing a happy monologue at Time 2 than Time 1. The increase in confidence may have manifested in certain facial cues of controls that were processed by observers and affected their judgments. The treatment-related increase in facial mobility of LSVT patients may have also facilitated the expression of increased comfort levels on their faces at Time 2. Conversely, the reduced facial mobility of the Untreated and ARTIC patients may have masked their feeling of ease and, as a result, prevented any changes in observers' Pers+ judgments. As the lack of significant increases on CV_H1, AU Variability, and AU Lability in the Control group suggests, these measures did not capture the facial cues signaling the increased comfort level of participants. As discussed in the next section regarding Hypothesis 2c, the lack of significant correlations between these facial measures and observers' judgments also suggests that other facial cues, not measured in this study, were the basis for the increased Pers+ ratings of LSVT and Control groups.

The present findings raise the question of why the Pers+ but not Beh+ ratings showed significant increases. We hypothesize that the reason for this difference is that the task of judging participants' behavior was implicitly interpreted by many observers as the requirement

to focus on only those cues that leave relatively little room for doubt. On the contrary, because personality is a less tangible construct than behavior, the task of judging participants' personality may have encouraged some observers to use more subtle cues as the basis for their judgments. Consequently, observers may have been more sensitive to changes in facial movement (from Time 1 to Time 2) when rating participants' personality, as compared to rating behavior.

The preliminary nature of the tests supporting Hypothesis 2b. The increase in the personality ratings of LSVT patients from Time 1 to Time 2 appears encouraging with regard to LSVT-related reduction of observers' negative impressions of PD patients. However, this finding needs to be interpreted cautiously because the effect of LSVT on the Pers+ ratings was evident only in additional analyses in which statistical significance levels were not adjusted for multiple comparisons. Also, the practical implications of this finding may be limited for two reasons. First, it is unknown whether the magnitude of this effect is sufficiently large to have a meaningful impact on the social interactions of the patient. It should be noted, however, that the increase of 0.21 in the Pers+ ratings of the LSVT group -- roughly 30% of the difference between the Time 1 Pers+ ratings of LSVT and Control groups (see Table 9) -- does appear to be encouraging. Second, similar to LSVT-related improvements in facial movement, any reductions in observers' negative impressions immediately after LSVT may not persist over time. As discussed below, we aim to collect the personality and behavior ratings of participants at a six-month follow-up to address this issue.

Hypothesis 2c: Facial Mobility as a Mediator of LSVT's Effect on Others' Perception of PD Patients

Investigations of facial cues as the basis of personality inferences have shown that such inferences are based on permanent facial features (Montepare & Zebrowitz, 1998; Zebrowitz,

1996, 1997) and transient facial expressions (Knutson, 1996; Marsh, Adams, & Kleck, 2005). We hoped to extend these findings by showing that such inferences are also based on the global measures of facial mobility (e.g., AU Variability). However, the correlations between PD patients' Time 1 scores on these measures and observers' Pers+ and Beh+ ratings of these patients were nonsignificant suggesting that personality and behavior inferences are made based on other aspects of appearance and/or behavior. Furthermore, Hypothesis 2c was not supported as none of the facial mobility measures used in this study mediated the effect of LSVT on the personality and behavior judgments of observers.

These negative findings raise the following question: what cues could have mediated LSVT-related increase in the Pers+ ratings? Because observers were instructed to base their ratings on participants' facial communication (see instructions to observers in the Appendix), we believe that facial cues that were not measured in this study mediated LSVT's effect on the Pers+ judgments. Based on previous studies implicating specific facial expressions as the basis for personality judgments (Marsh et al., 2005), we hypothesize that the frequency and/or intensity of specific AU combinations influenced observers' judgments and mediated the effect of LSVT. Further, because, due to increased facial mobility, LSVT patients probably expressed positive feelings more frequently and/or intensely during the happy monologues, the increased incidence (or intensity) of happy facial expressions may have acted as a mediator of LSVT's effect. A previous study showing that people with a happy facial expression are perceived as higher in extraversion (Borkenau & Liebler, 1992) than the same individuals with a neutral expression supports this interpretation. Because higher Pers+ means reflected increased scores on a number of positive attributes, including extraversion, this investigation predicts the observed increases in the Pers+ ratings of LSVT patients due to more frequent (or intense) happy expressions in this

group after than before treatment. According to this account, a higher frequency (or intensity) of happy expressions should be associated with higher Pers+ ratings. We plan to test this prediction in future studies.

Future Directions

What is the mechanism underlying the effect of LSVT on facial movement? LSVT is hypothesized to achieve therapeutic gains by normalizing the inadequate computation of vocal amplitude in the parkinsonian brain (i.e., a mechanism termed "rescaling" by LSVT proponents [Sapir, Ramig, & Fox, 2011]). If, as we hypothesize, LSVT affects facial movement through the same mechanism, it is likely that LSVT would increase the amplitude (i.e., reduce hypometria) of facial movement, resulting in greater AU intensity. Because the mechanisms of bradykinesia and hypometria overlap (Berardelli et al., 2001), LSVT is also expected to lessen the bradykinesia of facial movement in PD (Bowers et al., 2006a), resulting in the shortening of AU onset durations (i.e., duration between the onset of an AU and its apex).

The hypothesis that LSVT raises the amplitude of facial movement also suggests that AU Variability and AU Lability increases observed in this study are manifestations of this effect. To relate the increases on these variables to reduced hypometria, we hypothesized earlier that due to hypometria, AUs that would otherwise be produced with high intensity are produced with low intensity, and AUs that would otherwise be produced with low intensity are not produced. If a given person has the tendency to produce certain AUs (e.g., AU 23, also known as the "lip tightener") with low intensity, the advance of PD would eliminate the production of these AUs in that person. By reducing hypometria, LSVT would allow patients to produce these low-intensity AUs, thereby increasing the overall number of AUs (increased AU Lability) and the range of different AUs (increased AU Variability). This account predicts that for a given LSVT patient,

AUs ordinarily produced with low intensity would show greater LSVT-related increases in their frequency than AUs ordinarily produced with high intensity.

To summarize, the hypothesis that LSVT affects facial movement primarily by reducing hypometria and bradykinesia predicts that LSVT would lead to: a) greater average AU intensity, b) shorter AU onset durations, and c) greater increases in the frequency of AUs that are usually produced with low-intensity, relative to AUs usually produced with high-intensity, by a given person. By quantifying the facial movements of participants in this study on additional measures, such as the intensity and onset duration of each AU, we plan to test these predictions in follow-up work.

Which facial events serve as the basis of personality and behavior judgments?

Besides measuring facial movement on additional variables for a more thorough analysis of the current dataset, we plan to examine changes in the frequency of several facial events from Time 1 to Time 2. This analysis will test the hypothesis that unlike changes in the aggregate measures of facial mobility (e.g., AU Lability), changes in the frequency of facial events that usually communicate emotionally salient information[42] gave rise to the increase in the Pers+ ratings of LSVT group and, more generally, serve as the basis of personality and behavior judgments. Because participants produced happy monologues, facial events involving smiles, such as AU 12 and AU 6+12, appear to be good candidates for probing this hypothesis. Thus, as in the tests of Hypothesis 2c, the mediation analysis will be used to examine whether increases in the

---

42 For instance, a facial event consisting of AUs 6, 15, and 17 is usually interpreted as a reaction of sadness.

frequency of these facial events from Time 1 to Time 2 mediate the effect of LSVT on the Pers+ ratings. Given the evidence suggesting that PD patients produce a lower number of smiles with a check raise than do controls (Simons et al., 2004; Smith et al., 1996), it is possible that the frequency of AU 6+12 (smile with a cheek raise) increased to a greater extent than the frequency of AU 12 (smile without a cheek raise) as a result of LSVT. Further, because smiles perceived as "false" -- which frequently correspond to smiles without a check raise (Frank et al., 1993; Gosselin et al., 2002) -- are produced by PD patients more often than by controls (Pitcairn et al., 1990), it is conceivable that increases in the frequency of those two facial events have the opposite effects on the Pers+ ratings. Taken together, these findings suggest the intriguing possibility that an LSVT-related increase in the proportion of "sincere" smiles (AU 6+12) leads to the improved perception of LSVT patients' personalities by observers. We plan to explore this interesting scenario by recording the occurrences of these smiles in a follow-up study.

The distinction between "sincere" (AU 6+12) and "false" (AU 12) smiles exemplifies the greater voluntary control in the production of AUs in the lower face, such as AU 12, compared to AUs in the upper face, such as AU 6. Given the different neuronal substrates of voluntary and involuntary facial movement (Morecraft et al., 2004), the differential effect of LSVT on muscle contractions in the upper versus lower face would be informative with regard to the neurological mechanisms of LSVT. For this reason, we plan to examine the effects of LSVT separately for the two types of facial movements in a future investigation.

Is the reduction of facial movement deficits in LSVT patients maintained at a six-month follow-up? Although the obtained effects of LSVT on facial movement in PD are encouraging, it is possible that these effects are transient. If this is the case, it is impractical to use a treatment based on the principles of LSVT to treat facial movement deficits in PD. Given

the psychosocial impact of these deficits, discussed earlier, the importance of determining whether such treatment can achieve long-lasting increases in facial mobility is paramount. Evidence that improvements in vocal characteristics in PD are maintained for at least two years after LSVT completion (Ramig et al., 2001) is encouraging in this respect. If, as we hypothesized earlier, the same neuronal mechanisms are involved in scaling the amplitude of voice and facial movement, it is conceivable that LSVT-related improvements of facial movement abnormalities also persist over time. To examine this possibility, the happy monologues produced by participants at a six-month follow-up are currently being collected. We plan to analyze participants' facial movement in these monologues to better understand the long-term benefits of LSVT.

Do the findings of this study generalize to the expression of other emotions? The comparisons of PD-related impairments in facial emotional expression across different emotion types suggest that both posed and spontaneous expression of happiness and disgust are especially impaired in this clinical population (Madeley et al., 1995; Simons et al., 2003 & 2004; Smith et al., 1996). In contrast, another study suggests that facial masking in PD diminishes when PD patients speak about happy topics (Takahashi et al., 2010). This discrepancy may be due to the fact that, with one exception (Smith et al., 1996), all the analyses showing that PD patients are impaired in the facial communication of happiness, relative to other emotions, involved posed emotional expressions. By contrast, Takahashi and colleagues examined spontaneous emotional expression, the focus of this investigation. If, as the Takahashi et al. study suggests, PD patients are most impaired when speaking about negative or neutral topics, it is important to ascertain that LSVT can improve PD patients' communication in these situations as well. To this end, we plan to analyze the facial movements of the same participants as they produced angry and sad

monologues. Although these data were not analyzed in the current study, participants produced these monologues at Times 1 and 2 as part of a larger project.

Study Limitations

The design of this study has several limitations that need to be taken into account when interpreting its results. First, although, as noted earlier (see the first subsection of the Discussion), the somewhat low overall intercoder reliability of .71 is unlikely to account for any of the present results, a higher reliability index would increase one's confidence in the study findings. Second, because participants' personalities were not assessed, it is unknown the extent to which genuine personality differences between PD patients and controls, as opposed to the negative bias in observers' perception of PD patients, contributed to the differences in how these two groups were perceived by observers at Time 1. Third, because observers in this study rated the personality and behavior of strangers based on the silent video clips, some of the present findings may not apply to other, more ecologically valid settings. For instance, it is unclear whether people familiar with a PD patient (e.g., the patient's physician or relatives) perceive the patient with a negative bias. As most human interactions occur among people familiar with one another, investigating the impact of PD outside the laboratory remains an important goal for future studies. Fourth, the personality and behavior rating forms included only a few dimensions used in previous studies (e.g., Pentland et al., 1987) showing that PD participants were judged negatively by others. Consequently, the present findings offer a limited and possibly biased description of how PD patients are perceived by others and how others' perceptions of these patients are impacted by LSVT. For instance, it may be that although LSVT patients were viewed more positively after, relative to before, treatment on the personality dimensions used in this study, observers' perception of these patients changed in the opposite direction on other

personality characteristics. This does not appear likely, as other studies (e.g., Tickle-Degnen et al., 2004) did not find any evidence for perceptual biases in the perception of PD participants on those other characteristics (i.e., the ones that were not included in this study). Nevertheless, it may be worthwhile to examine observers' perceptions of PD participants in a more comprehensive fashion. Fifth, because participants did not complete the BDI at Time 2, it is possible that some participants became clinically depressed. Because clinically depressed individuals, relative to healthy controls, show fewer facial emotional expressions and less facial muscle activity while viewing or imagining positively valenced stimuli (Gehricke & Shapiro, 2000; Katsikitis & Pilowsky, 1991), the low AU Variability and AU Lability scores of some participants at Time 2 may have resulted from their depression rather than PD. Thus, intergroup differences in the proportion of individuals who were clinically depressed at Time 2 may have spuriously caused significant intergroup differences on AU Variability Δ and AU Lability Δ and thus confounded the effect of LSVT on facial movement. Although the short period of only one month between the first and second assessments makes it unlikely that a substantial number of participants were clinically depressed at Time 2, knowing participants' BDI scores at Time 2 would have made the present findings more convincing. Seventh, information regarding the fidelity of ARTIC and LSVT. Due to the lack of this information, it is unknown whether the two clinicians administering these treatments deviated from the standardized treatment protocol (Ramig, Pawlas, & Countryman, 1995) or whether the patients completed their homework assignments. Although the effort to make both treatments identical on a number of variables (e.g., treatment intensity, see the Method section) makes it unlikely that the present findings resulted from patient noncompliance or clinicians' deviations from the manual, evidence for the fidelity of LSVT and ARTIC administration would have strengthened the study conclusions.

Finally, observers in this study were drawn from a convenience sample of young students from courses in introductory psychology. Thus the findings regarding their perception of participants may not be generalizable to other age groups. Though negative biases in the perception of PD patients were also observed among their age peers (Takahashi et al., 2010), future studies should attempt to recruit a random sample of observers.

Summary of conclusions

The present study demonstrated that during the production of happy monologues, PD patients, compared to demographically matched controls, showed a decreased repertoire (AU Variability) and rate of change (AU Lability) of facial expressions. The available evidence suggests that these impairments of facial movement may disrupt the lives of PD patients by causing a misdiagnosis of depression, delaying the treatment of depression in PD patients and worsening their cognitive decline, and leading to a negative bias in the perception of PD patients by others. Remarkably, despite this negative bias, observers appear to be sensitive to PD patients' symptoms of subclinical depression. Besides facial movement deficits, a cluster of negative personality characteristics, which appears to be more common among PD patients than others, may contribute to the negative perception of PD patients.

Study results showed that LSVT, relative to control articulation therapy, can alleviate the facial movement deficits of PD patients. The available evidence from other studies suggests at least three reasons for these effects: LSVT is intensive and directly targets the deficits of PD patients, LSVT is parameter-specific rather than task-specific, and the contractions of the lower face muscles are closely coordinated with vocal amplitude, the target of LSVT. The present study provided preliminary evidence that LSVT can also improve others' perception of PD patients. We conclude that LSVT-related increases in the frequency and/or intensity of specific

AU combinations, rather than increases in overall facial mobility, mediated LSVT's effect on observers' perception of PD patients.

Tables

Table 1
Single AUs in the Facial Action Coding System.

AU number	FACS Name	Muscular Basis
1.	Inner Brow Raiser	Frontalis, Pars Medialis
2.	Outer Brow Raiser	Frontalis, Pars Lateralis
4.	Brow Lowerer	Depressor Glabellae, Depressor Supercilli; Corrugator
5.	Upper Lid Raiser	Levator Palpebrae Superioris
6.	Cheek Raiser	Orbicularis Oculi, Pars Orbitalis
7.	Lid Tightener	Orbicularis Oculi, Pars Palebralis
9.	Nose Wrinkler	Levator Labii Superioris, Alaeque Nasi
10.	Upper Lip Raiser	Levator Labii Superioris, Caput Infraorbitalis
11.	Nasolabial Fold Deepener	Zygomatic Minor
12.	Lip Corner Puller	Zygomatic Major
13.	Cheek Puffer	Caninis
14.	Dimpler	Buccinator
15.	Lip Corner Depressor	Triangularis
16.	Lower Lip Depressor	Depressor Labii
17.	Chin Raiser	Mentalis
18.	Lip Puckerer	Incisivii Labii Superioris; Incisivii Labii Inferioris
20.	Lip Stretcher	Risorius
22.	Lip Funneler	Orbicularis Oris
23.	Lip Tightener	Orbicularis Oris
24.	Lip Pressor	Orbicularis Oris
25.	Lips Part	Depressor Labii, or Relaxation of Mentalis or Orbicularis Oris
26.	Jaw Drop	Masetter; Temporal and Internal Pterygoid Relaxed
27.	Mouth Stretch	Pterygoids; Digastric
28.	Lips Suck	Orbicularis Oris

Note. From “Introduction: The study of spontaneous facial expressions in psychology,” by E. L. Rosenberg, In P. Ekman, & E. L. Rosenberg (Eds.), What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS) (2nd ed.; pp. 14). New York, NY: Oxford University Press.

Table 2
More grossly defined AUs in the Facial Action Coding System.

AU number	FACS Name
8.	Lips Toward Each Other
19.	Tongue Out
21.	Neck Tightener
29.	Jaw Thrust
30.	Jaw Sideways
31.	Jaw Clencher
32.	Lip Bite
33.	Blow
34.	Puff
35.	Cheek Suck
36.	Tongue Bulge
37.	Lip Wipe
38.	Nostril Dilator
39.	Nostril Compressor
43.	Eyes Closure
45.	Blink
46.	Wink

Note. From “Introduction: The study of spontaneous facial expressions in psychology,” by E. L. Rosenberg, In P. Ekman, & E. L. Rosenberg (Eds.), What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS) (2nd ed.; pp. 15). New York, NY: Oxford University Press.

Table 3
Participant Demographics and Relevant Medical Variables

Relevant variable		Participant Group				
		ARTIC	LSVT	Untreated	Control	Total
Sample Size		12	16	17	11	56
Age		69.3 (10.3)[1]	68.5 (6.7)	65.7 (8.9)	61.8 (8.6)	66.5 (8.8)
Gender						
	Male	8 (66.7 %)	12 (75.0%)	13 (76.5%)	4 (36.4%)	37 (66%)
	Female	4 (33.3%)	4 (25.0%)	4 (23.5%)	7 (63.6%)	19 (34%)
Race/ Ethnicity						
	Caucasian	12 (100%)	12 (75%)	16 (94%)	10 (91%)	50 (89%)
	Hispanic	0 (0%)	2 (13%)	1 (6%)	0 (0%)	3 (5%)
	Asian	0 (0%)	1 (6%)	0 (0%)	0 (0%)	1 (2%)
	African American	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
	Other	0 (0%)	1 (6%)	0 (0%)	1 (9%)	2 (4%)
Years of education		16.0 (3.7)	15.6 (2.7)	15.5 (3.1)	17 (2.2)	15.9 (3.0)
Hoehn-Yahr Stage		2.2 (0.7)	2.2 (0.5)	2.1 (0.6)	N/A	2.2 (0.6)[2]
Years since diagnosis		5.1 (4.0)	5.9 (7.1)	6.6 (5.9)	N/A	6.0 (5.8)[2]
# of PD medications		1.7 (1.3)	2.0 (0.9)	2.3 (1.1)	N/A	2.0 (1.1)[2]
BDI score at Time 1		8.6 (6.1)	10.1 (5.9)	8.4 (5.8)	2.3 (2.4)	7.7 (6.0)

[1] For variables Gender and Race/Ethnicity, this table shows frequency counts and percentages (shown in parentheses). For all other variables, the mean and standard deviation (shown in parentheses) are presented.
[2] These values are given for the total sample of PD patients (n = 45).

Table 4
Observer Demographics and Relevant Medical Variables

Relevant variable		Set of video clips				
		Set I	Set II	Set III	Set IV	Total
Number of observers		26	29	28	27	110
Age		19.5 (1.4)[1]	19.7 (1.7)	19.6 (1.5)	19.4 (1.3)	19.6 (1.5)
Gender						
	Male	10 (38.5 %)	12 (41.4%)	11 (39.3%)	12 (44.4%)	45 (40.9%)
	Female	16 (61.5%)	17 (58.6%)	17 (60.7%)	15 (55.6%)	65 (59.1%)
Race/ Ethnicity						
	Caucasian	14 (53.8%)	14 (48.3%)	14 (50.0%)	15 (55.6%)	57 (51.8%)
	Hispanic	4 (15.4%)	6 (20.7%)	4 (14.3%)	2 (7.4%)	16 (14.5%)
	Asian	1 (3.8%)	4 (13.8%)	3 (10.7%)	6 (22.2%)	14 (12.7%)
	African American	1 (3.8%)	2 (6.9%)	2 (7.1%)	0 (0.0%)	5 (4.5%)
	Other	6 (23.1%)	3 (10.3%)	5 (17.9%)	4 (14.8%)	18 (16.4%)
Years of education		13.6 (1.2)	13.4 (1.3)	13.7 (1.2)	13.8 (1.4)	13.6 (1.3)

[1] For variables Gender and Race/Ethnicity, this table shows frequency counts and percentages (shown in parentheses). For all other variables, the mean and standard deviation (shown in parentheses) are presented.

Table 5
Pearson Correlations among Facial Movement Measures at Time 1 for PD Patients, Normal Controls, and Total Participant Sample

Measure	1	2	3	Mean	Standard deviation
PD patients (n = 45)					
1. AU Variability	-	.81	.67	3.8	2.79
2. AU Lability		-	.58	7.27	7.35
3. AU Complexity			-	1.51	0.79
Normal controls (n = 11)					
1. AU Variability	-	.97	.48	15.27	10.98
2. AU Lability		-	.40	7	4.77
3. AU Complexity			-	1.83	0.97
Total sample (n=56)					
1. AU Variability	-	.84	.42	4.43	3.47
2. AU Lability		-	.38	8.84	8.69
3. AU Complexity			-	1.57	0.83

Table 6
Pearson Correlations among Changes (from Time 1 to Time 2) in Facial Movement Measures for Each Participant Group and Total Sample

Measure	1	2	3	Mean	Standard deviation
			ARTIC		
1. AU Variability Δ	-	.91	.43	-0.58	3.15
2. AU Lability Δ		-	.39	-1.25	6.68
3. AU Complexity Δ			-	0.11	0.75
			LSVT		
1. AU Variability Δ	-	.88	.49	1.81	3.04
2. AU Lability Δ		-	.56	5.00	8.24
3. AU Complexity Δ			-	0.40	1.31
			Untreated		
1. AU Variability Δ	-	.81	.38	-0.94	2.33
2. AU Lability Δ		-	.19	-1.71	6.82
3. AU Complexity Δ			-	-0.09	0.60
			Normal controls		
1. AU Variability Δ	-	.70	.38	-0.73	2.65
2. AU Lability Δ		-	.32	-1.18	7.77
3. AU Complexity Δ			-	0.47	0.73
			Total sample		
1. AU Variability Δ	-	.84	.42	0.20	0.91
2. AU Lability Δ		-	.38	0.41	7.78
3. AU Complexity Δ			-	-0.04	2.96

Table 7
Means of Three Original Facial Movement Measures and CV_H1 for Each Participant Group

	AU Variability		AU Lability		AU Complexity		CV_H1[1]
Group	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	
ARTIC	3.83 (3.76)[1]	3.25 (2.00)	6.58 (7.04)	5.33 (3.82)	1.27 (0.57)	1.38 (0.76)	-0.21 (1.09)
LSVT	3.38 (2.60)	5.19 (3.35)	6.68 (6.72)	11.69 (9.76)	1.45 (0.99)	1.85 (1.02)	0.68 (1.08)
Untreated	4.18 (2.24)	3.24 (1.71)	8.29 (8.39)	6.59 (5.01)	1.74 (0.68)	1.65 (0.67)	-0.30 (0.87)
Normal controls	7.00 (4.77)	6.27 (3.64)	15.27 (10.98)	14.09 (7.30)	1.83 (0.97)	2.31 (0.94)	-0.30 (0.95)
Total	4.5 (3.31)	4.39 (2.75)	8.84 (8.28)	9.25 (6.96)	1.57 (0.82)	1.78 (0.86)	0.00 (1.06)

[1] CV_H1 is a canonical variate created for testing Hypothesis 1. See text for details.
[2] Standard deviations are shown in parentheses.

Table 8
Covariance of Contrast 2: the CV_H1 Scores of ARTIC Versus LSVT Patients

Covariate	df	t	p
Participant's gender	26	1.63	.12
Participant's age	26	1.89	.07
Participant's handedness	26	1.97	.06
Hoehn-Yahr stage	26	1.85	.08
Time since PD diagnosis	26	1.80	.08
Participant's education level	26	1.85	.08
BDI score at Time 1	26	1.74	.09
MMSE	26	1.87	.07
Number of PD medications taken	26	2.00	.06

Table 9
The Means of the Overall+, Pers+, and Beh+ Ratings of Each Participant Group

	Pers+		Beh+		Overall+	
Group	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
ARTIC	4.17 (0.78)[1]	4.15 (0.66)	4.16 (0.87)	4.32 (0.72)	4.17 (0.73)	4.23 (0.60)
LSVT	3.87 (0.57)	4.08 (0.68)	3.83 (0.70)	3.91 (1.03)	3.85 (0.54)	4.03 (0.66)
Untreated	4.21 (0.61)	4.19 (0.61)	4.41 (0.66)	4.51 (0.63)	4.32 (0.54)	4.35 (0.55)
Normal controls	4.55 (0.97)	4.97 (0.93)	4.94 (1.02)	5.20 (0.97)	4.74 (0.93)	5.08 (0.83)
PD patients (ARTIC, LSVT, and Untreated groups combined)	4.09 (0.67)	4.14 (0.65)	4.13 (0.78)	4.24 (0.85)	4.11 (0.64)	4.20 (0.62)
Total	4.20 (0.79)	4.34 (0.81)	4.33 (0.92)	4.48 (0.97)	4.27 (0.77)	4.42 (0.78)

[1] Standard deviations are shown in parentheses.

Table 10
The Means Compared in Contrasts 1 and 2 and the Corresponding Statistics

	ARTIC		LSVT		Non-LSVT participants		WJ statistic	p-value
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2		
				Pers+				
Contrast 1	-	-	3.87 (0.57)[1]	4.08 (0.68)	4.31 (0.82)	4.43 (0.83)	0.61	.43
Contrast 2	4.17 (0.78)	4.15 (0.66)	3.87 (0.57)	4.08 (0.68)	-	-	3.07	.08
				Beh+				
Contrast 1	-	-	3.83 (0.70)	3.91 (1.03)	4.50 (0.92)	4.68 (0.87)	0.50	.48
Contrast 2	4.16 (0.87)	4.32 (0.72)	3.83 (0.70)	3.91 (1.03)	-	-	0.24	.62
				Overall+				
Contrast 1	-	-	3.85 (0.54)	4.03 (0.66)	4.41 (0.79)	4.55 (0.77)	0.15	.69
Contrast 2	4.17 (0.73)	4.23 (0.60)	3.85 (0.54)	4.03 (0.66)	-	-	1.06	.30

[1] Standard deviations are shown in parentheses.

Table 11
Mean Ratings of PD Patients and Normal Controls on Each Personality and Behavior Scale

Scale	Mean		WJ statistic	p-value	Effect size
	PD patients	Normal controls			
Personality scales					
Extraverted - Introverted	3.70 (1.07)[1]	4.80 (1.54)	49.18	< .01	0.88
Secure - Anxious	4.03 (0.97)	4.45 (1.29)	9.74	< .01	0.46
Independent - Dependent	4.33 (0.97)	4.23 (1.23)	0.70	.41	0.06
Behavior scales					
Engaged - Bored	4.64 (1.27)	5.48 (1.24)	38.05	< .01	0.63
Relaxed - Tense	3.71 (1.12)	4.62 (1.35)	39.89	< .01	0.81
Happy - Sad	3.72 (1.03)	4.58 (1.28)	40.19	< .01	0.82
Openhearted - Guarded	4.30 (1.17)	5.08 (1.32)	39.89	< .01	0.66
Liked – Disliked (averaged across two scales)	4.30 (0.69)	4.82 (0.82)	36.08	< .01	0.71

[1] Standard deviations are shown in parentheses.

Table 12
Pearson Correlations of BDI Score at Time 1 with Pers+, Beh+, and Overall+ Ratings for PD patients, Normal Controls, and Total Participant Sample

	PD patients		Normal controls		Total participant sample	
Measure	r	p	r	p	r	p
Pers+	-.45	< .01	.55	.08	-.38	< .01
Beh+	-.37	.01	.57	.07	-.38	< .01
Overall+	-.41	< .01	.57	.07	-.39	< .01

Table 13
The Means of the Overall+, Pers+, and Beh+ Ratings of 45 PD Patients

	Time 1			Time 2		
	Pers+	Beh+	Overall+	Pers+	Beh+	Overall+
PD patients with low BDI scores (n = 34; BDI ≤ 13)	4.21 (0.60)[1]	4.28 (0.77)	4.25 (0.68)	4.34 (0.63)[1]	4.47 (0.73)	4.41 (0.67)
PD patients with moderate BDI scores (n = 13; BDI > 13)	3.72 (0.49)	3.74 (0.42)	3.73 (0.42)	3.82 (0.48)	3.98 (0.42)	3.91 (0.43)
Total	4.09 (0.61)	4.15 (0.74)	4.12 (0.66)	4.27 (0.99)	4.35 (0.70)	4.29 (0.65)

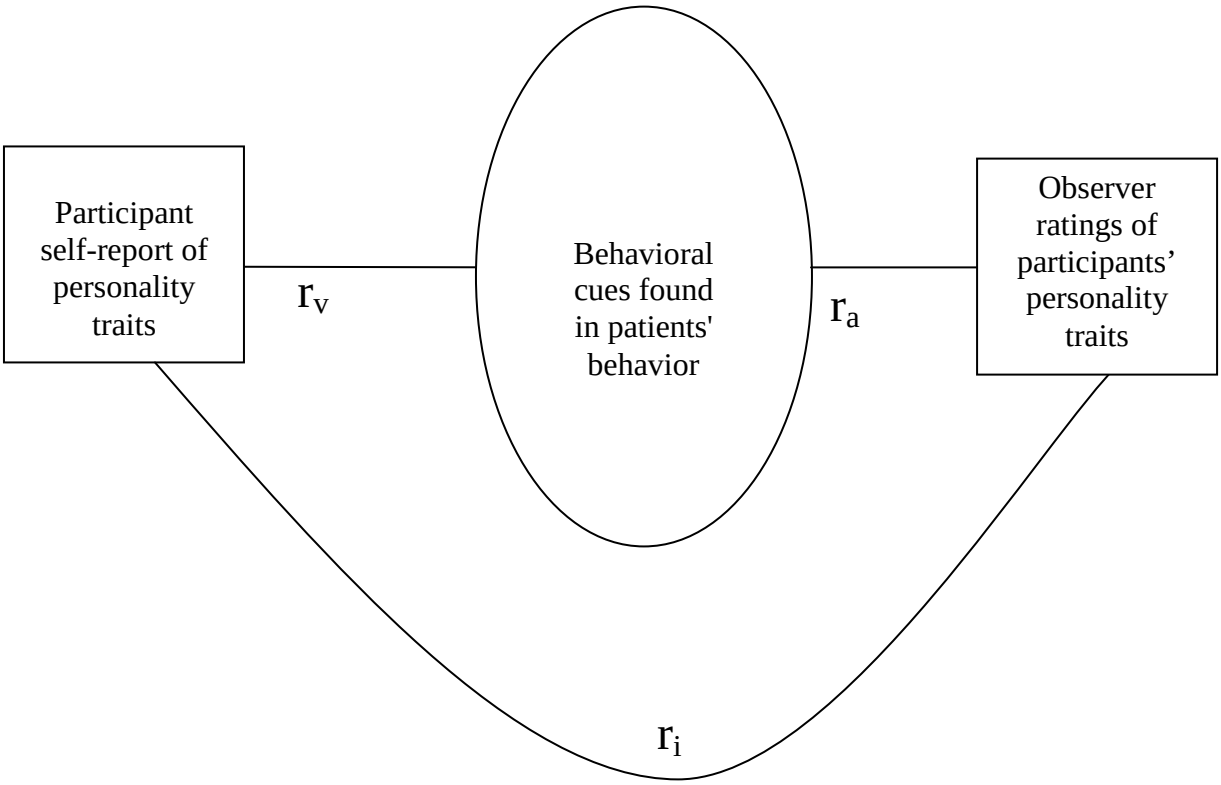
[1] Standard deviations are shown in parentheses.

Table 14
The Means of the Overall+, Pers+, and Beh+ Ratings of 18 PD Patients (with BDI < 7) and Normal Controls

	Pers+	Beh+	Overall+
PD patients	4.17 (0.97)[1]	4.15 (1.01)	4.16 (0.84)
Normal controls	4.55 (0.97)	4.94 (1.02)	4.74 (0.93)
Total	4.27 (0.99)	4.36 (1.07)	4.31 (0.90)

[1] Standard deviations are shown in parentheses.

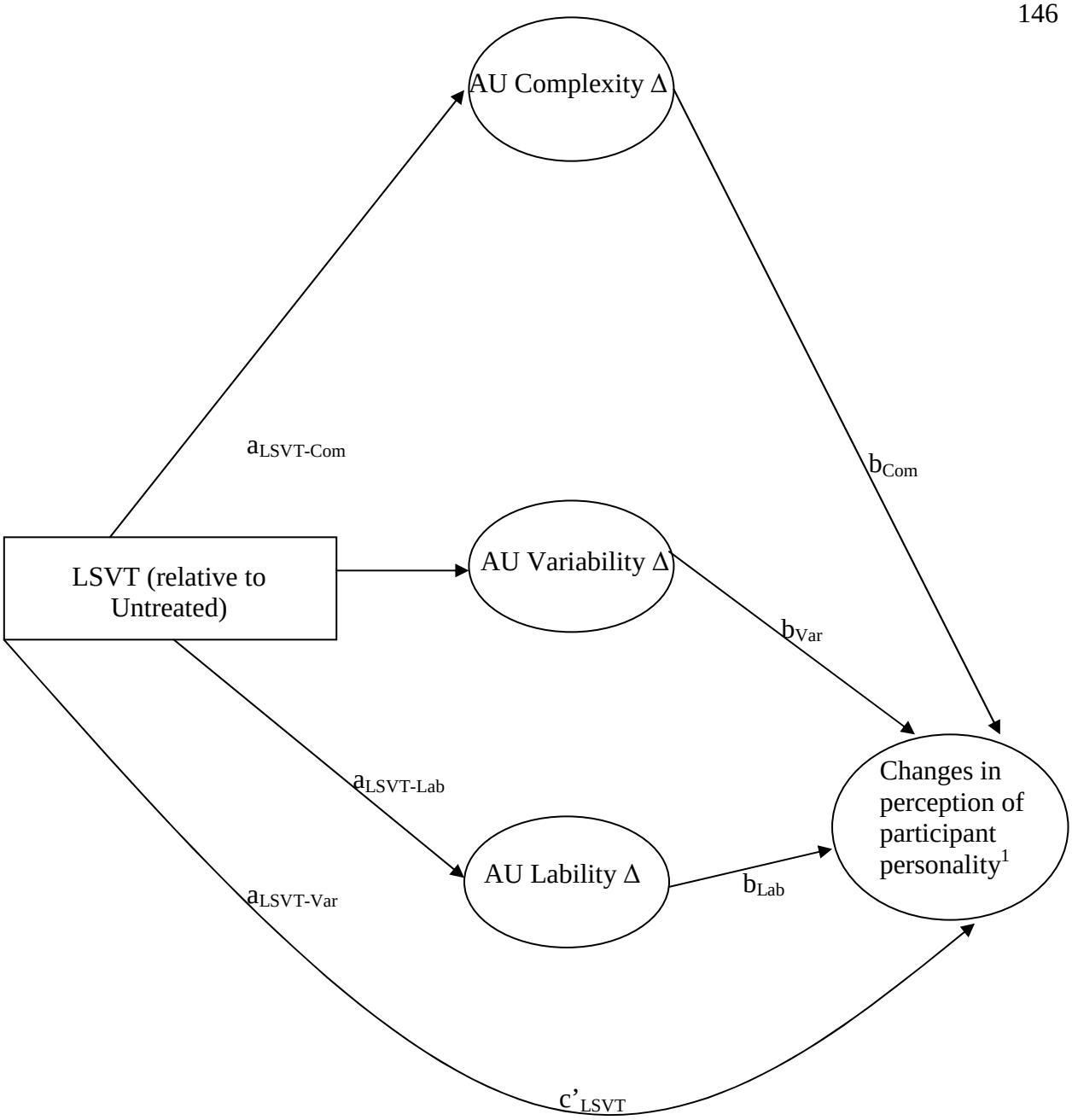
Figures



r_v - correlation between the patients' scores on a personality attribute and the presence of a behavioral cue (its frequency, intensity, or duration in the patients' behavior); r_i - correlation between the presence of a behavioral cue and the observers' ratings of a personality attribute in the patients; r_a - correlation between the patients' scores and the observers' ratings on a personality attribute. See text for explanation.

Note. From "Inferring personality traits of clients with Parkinson's disease from their descriptions of favourite activities," by K. D. Lyons, L. Tickle-Degnen, and E. J. DeGroat, 2005, *Clinical Rehabilitation, 19*(7), p. 801.

Figure 1. Application of Brunswik's lens model to research on the role of nonverbal cues in the perception of PD patients.



[1] Changes in the perception of participant personality will be measured as a change in ratings. See text for explanation.

Figure 2. The hypothesized mediation model of LSVT's effect on the perceived personality and behavior differences.

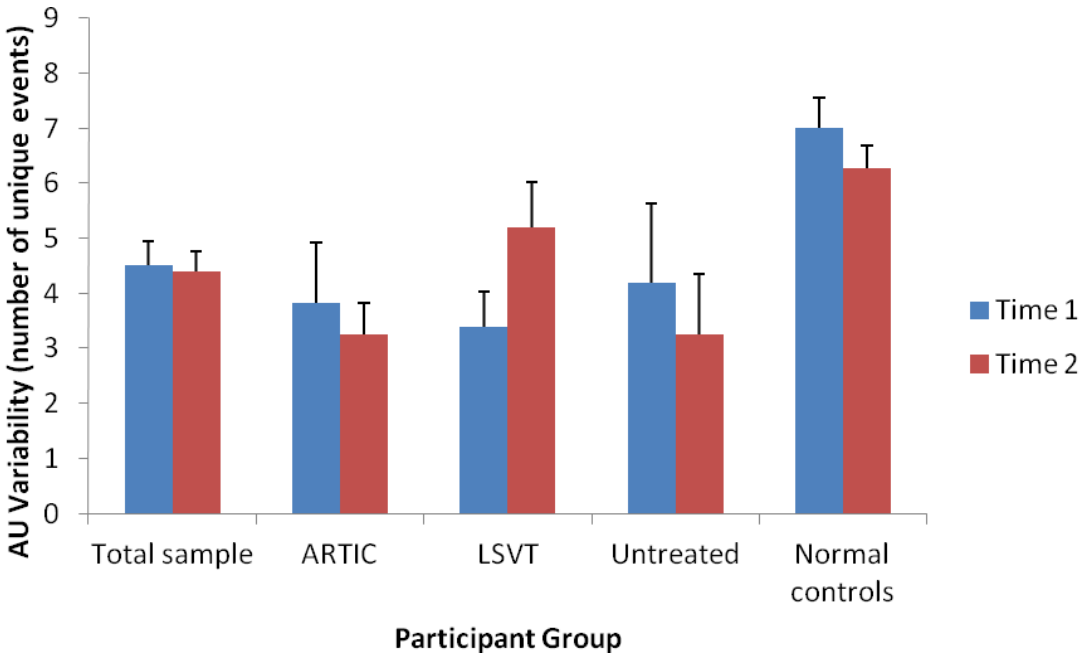


Figure 3. AU Variability means (+SE) at Times 1 and 2 by participant group.

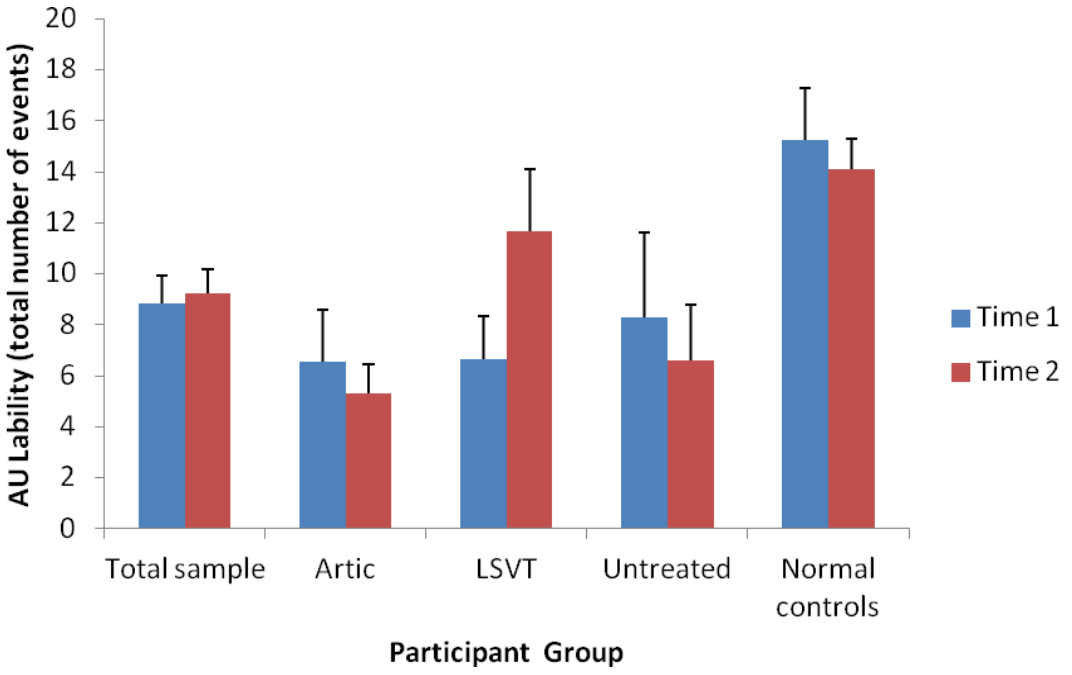


Figure 4. AU Lability means (+SE) at Times 1 and 2 by participant group.

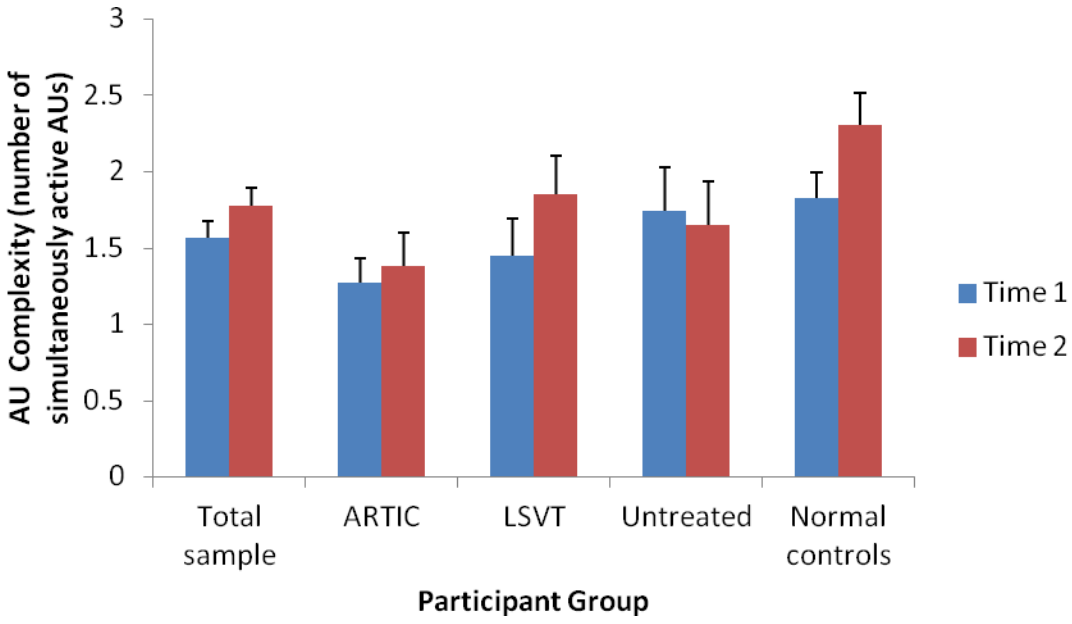


Figure 5. AU Complexity means (+SE) at Times 1 and 2 by participant group.

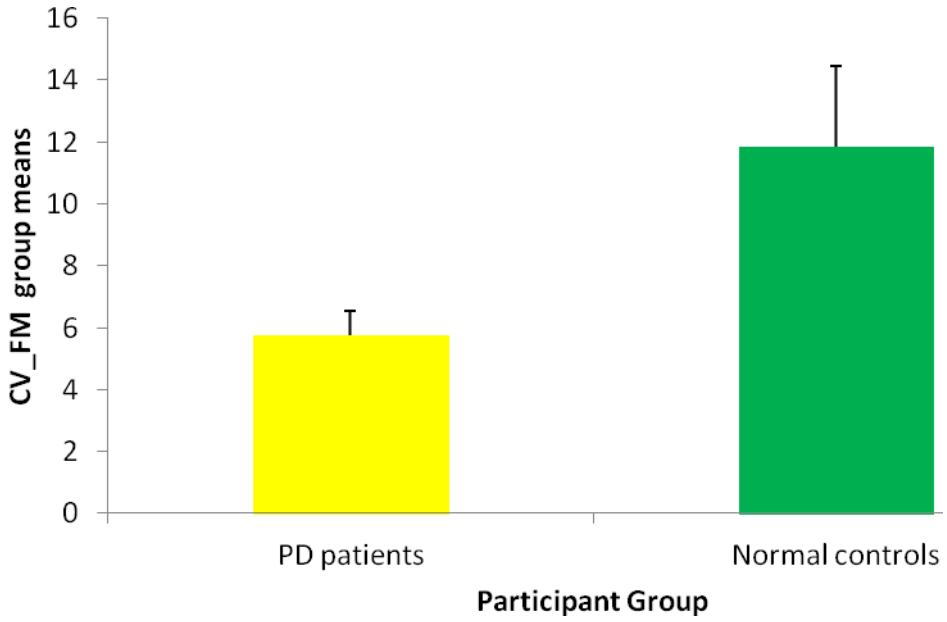


Figure 6. CV_FM means (+SE) for PD patients (n = 45) versus normal controls (n = 11).

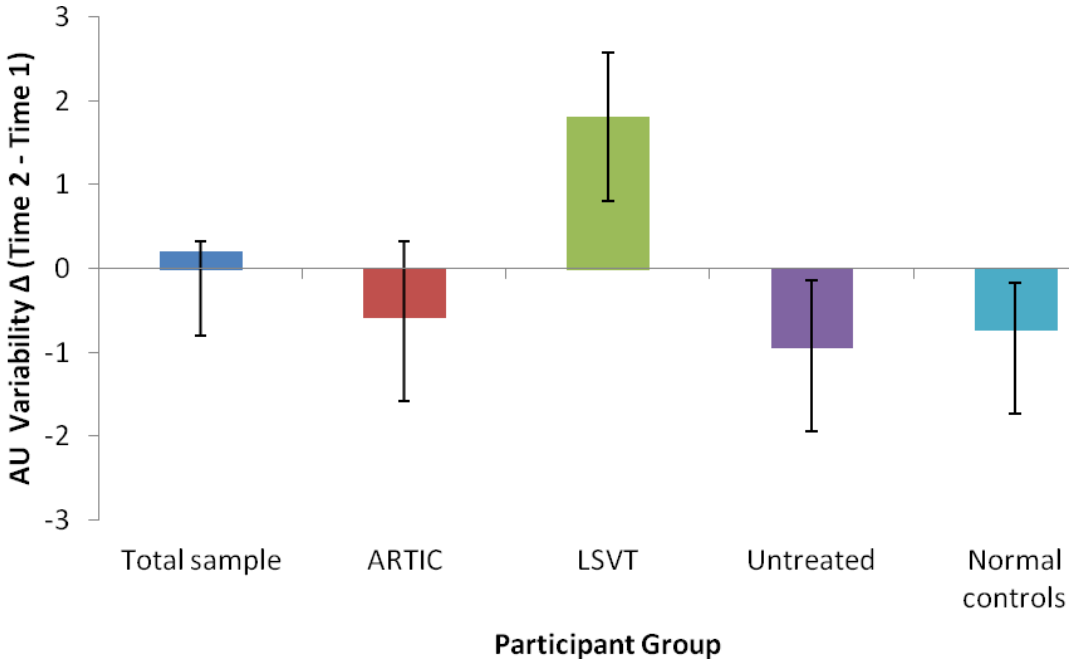


Figure 7. AU Variability Δ means (+SE) by participant group.

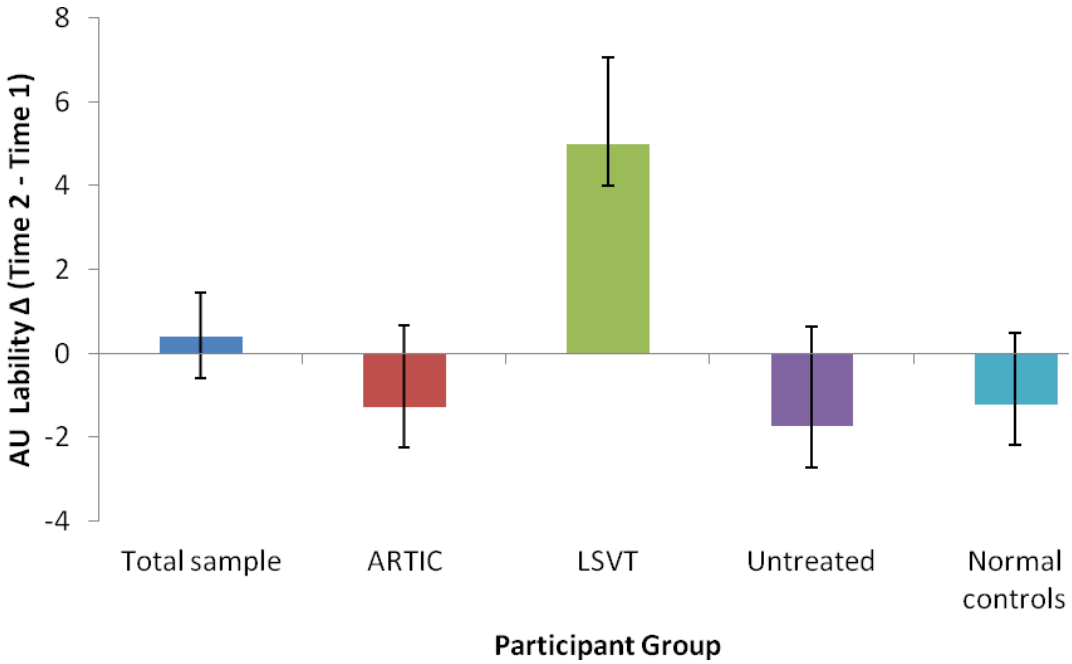


Figure 8. AU Lability Δ means (+SE) by participant group.

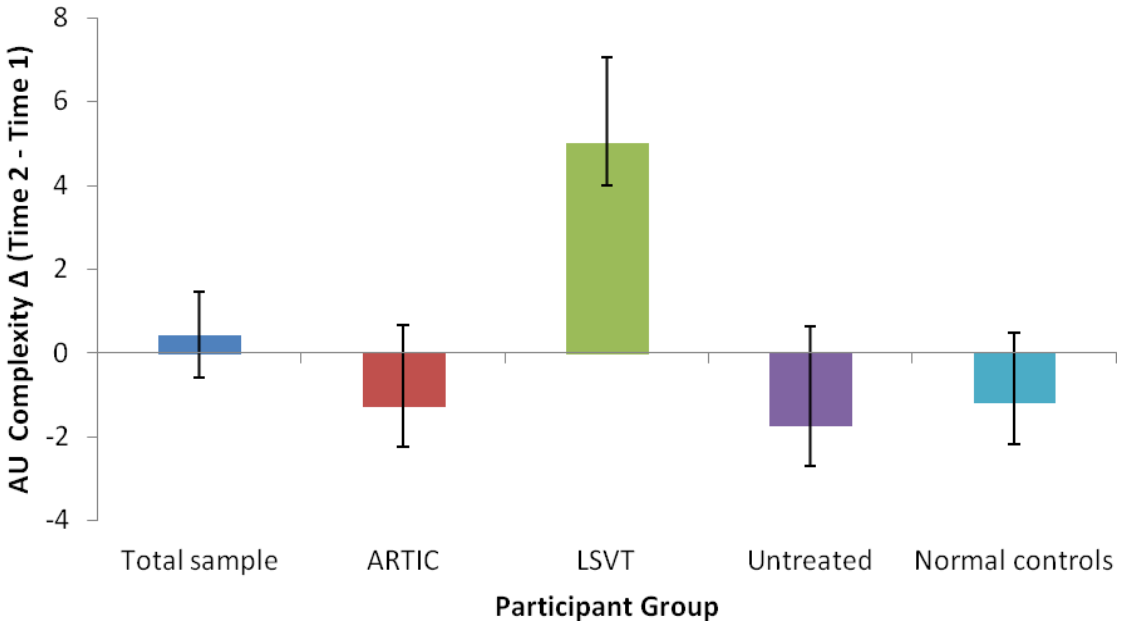


Figure 9. AU Complexity Δ means (+SE) by participant group.

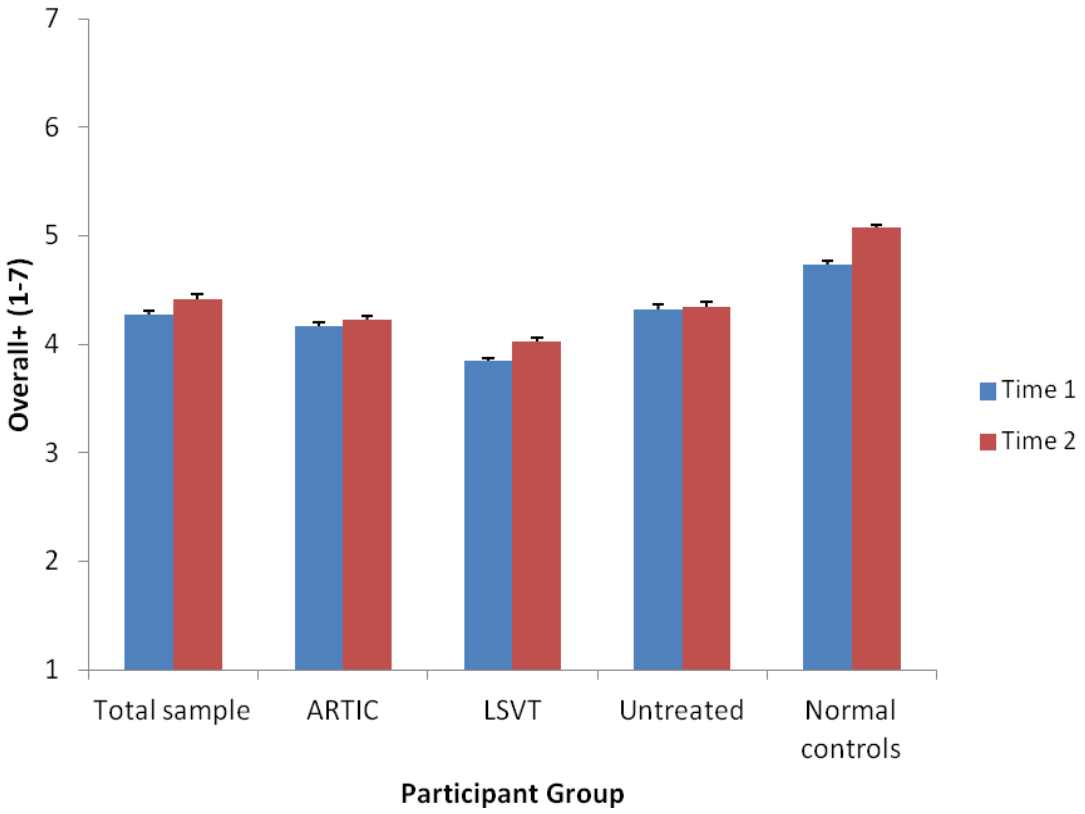


Figure 10. Overall+ means (+SE) at Times 1 and 2 by participant group.

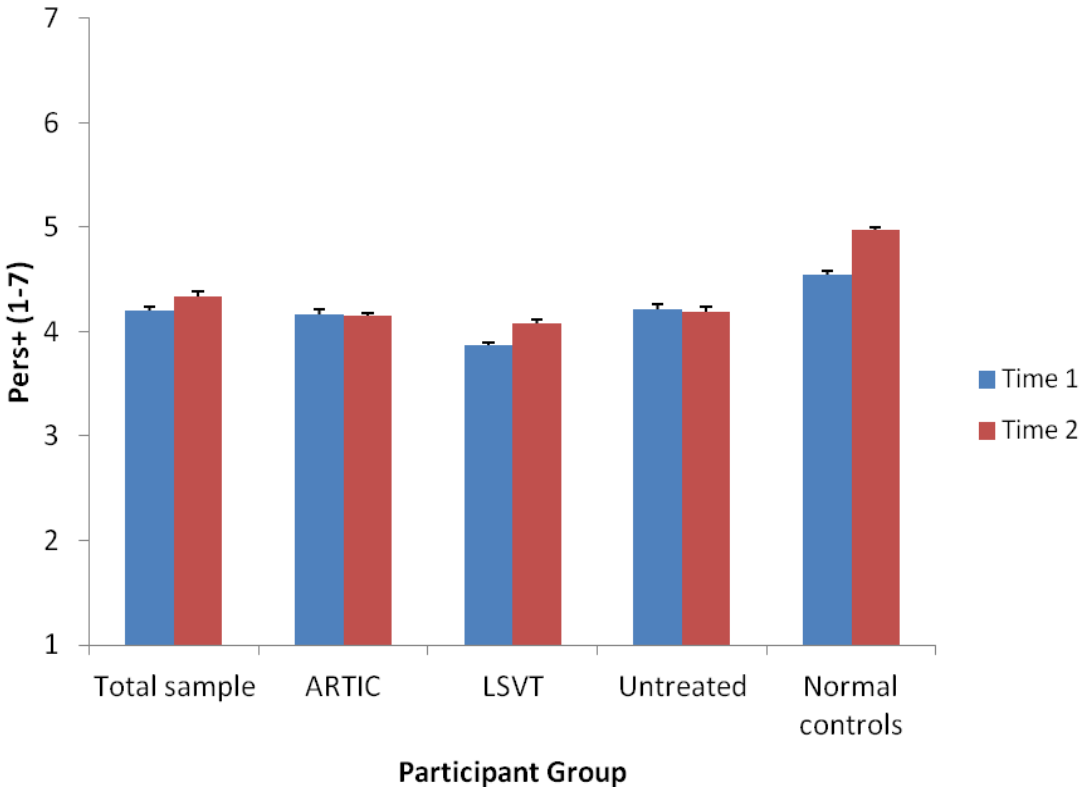


Figure 11. Pers+ means (+SE) at Times 1 and 2 by participant group.

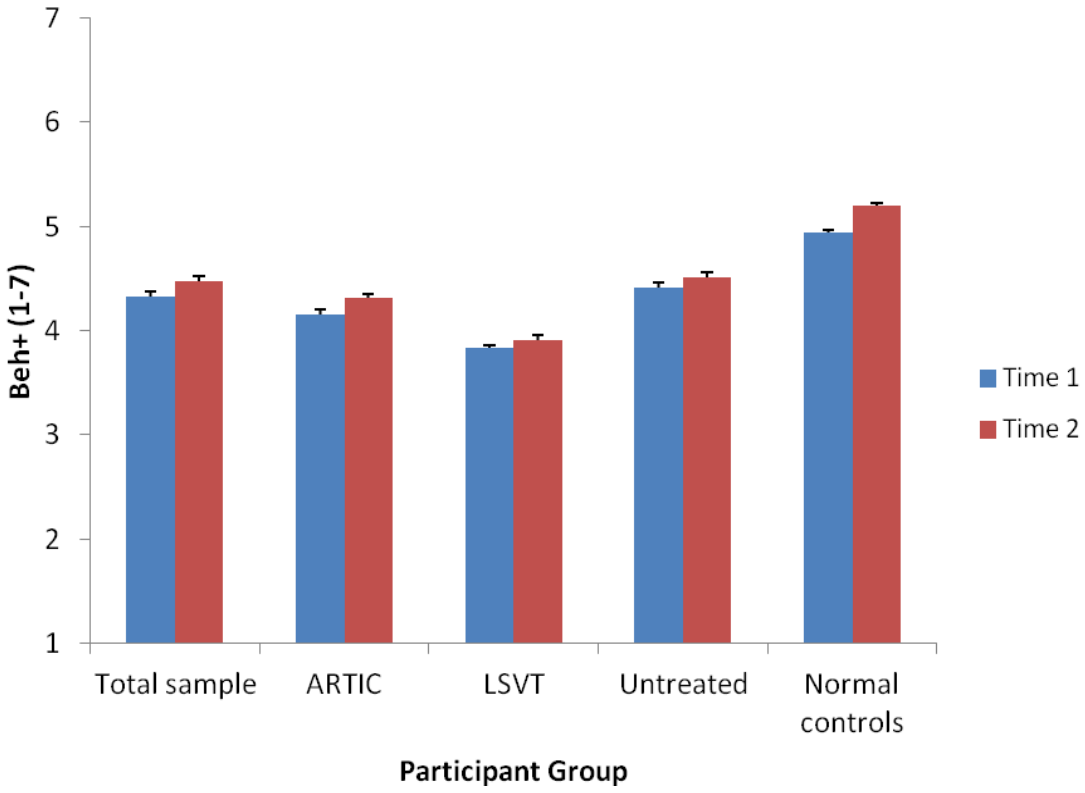


Figure 12. Beh+ means (+SE) at Times 1 and 2 by participant group.

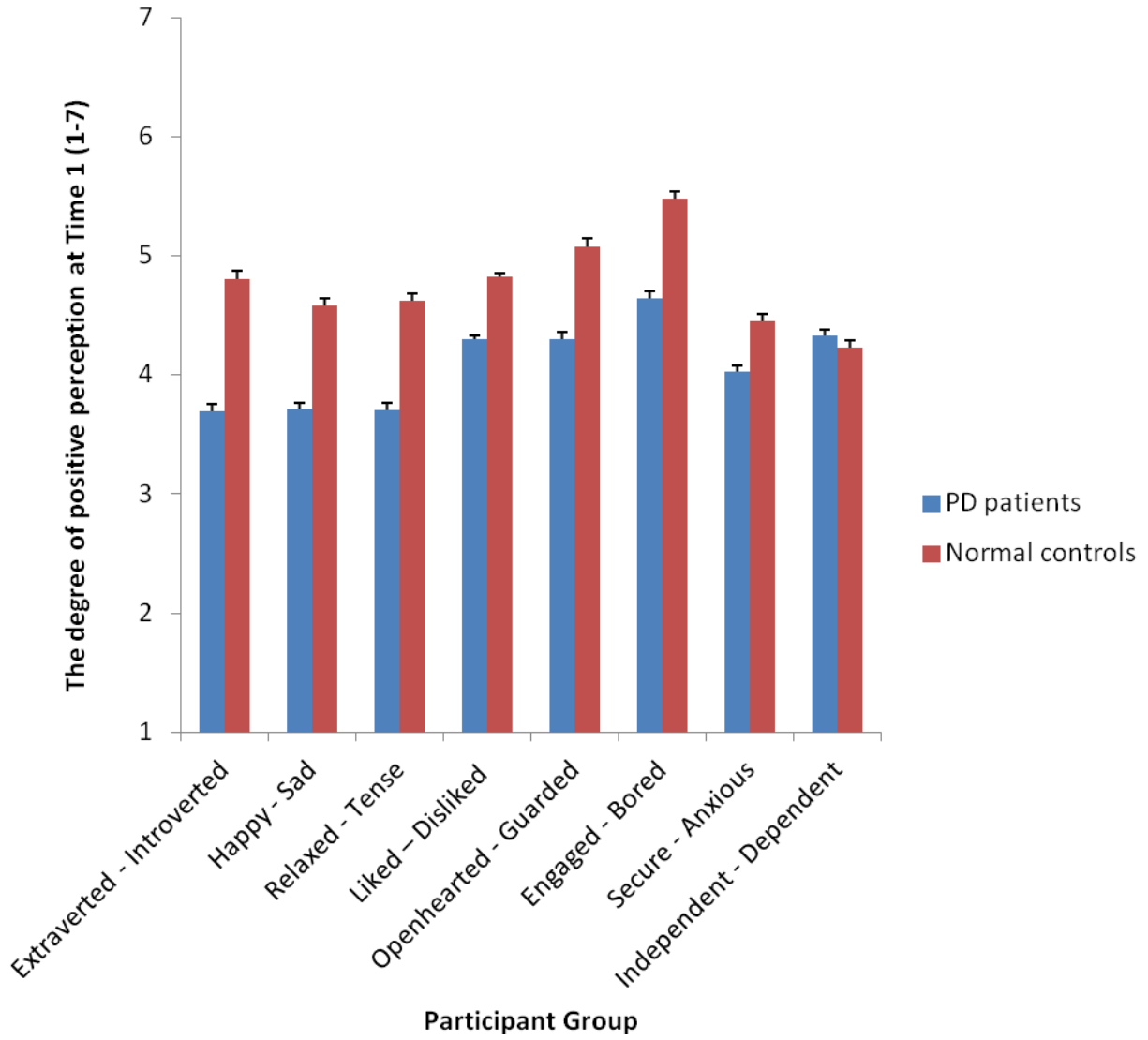


Figure 13. Time 1 mean ratings of PD patients and normal controls on each personality and behavior scale. The traits are ordered on the X-axis by the magnitude of the corresponding effect size (i.e., the effect size for the difference between PD patients and normal controls on that trait; see Table 9), with the traits associated with greater effect sizes presented on the left.

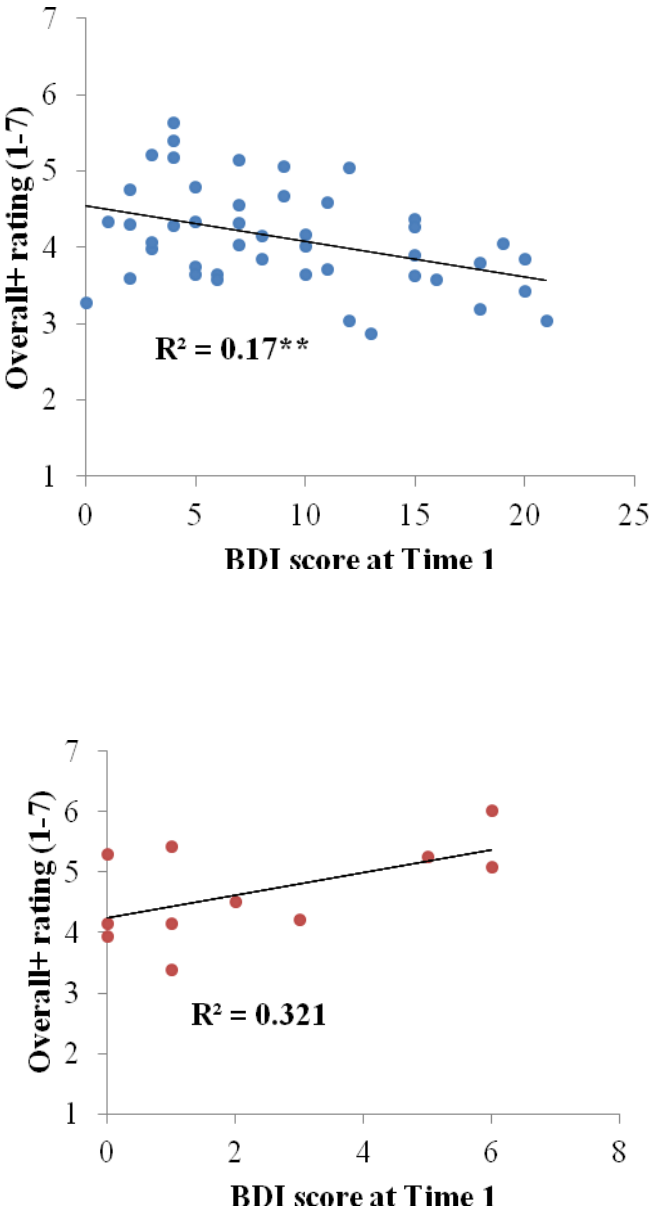


Figure 14. Relationship between BDI scores at Time 1 and Overall+ ratings at Time 1 for PD patients (top) and normal controls (bottom). Higher BDI scores reflect a greater amount of depressive symptomatology. Double asterisk indicates that a relationship was significant at the $p < .01$ level.

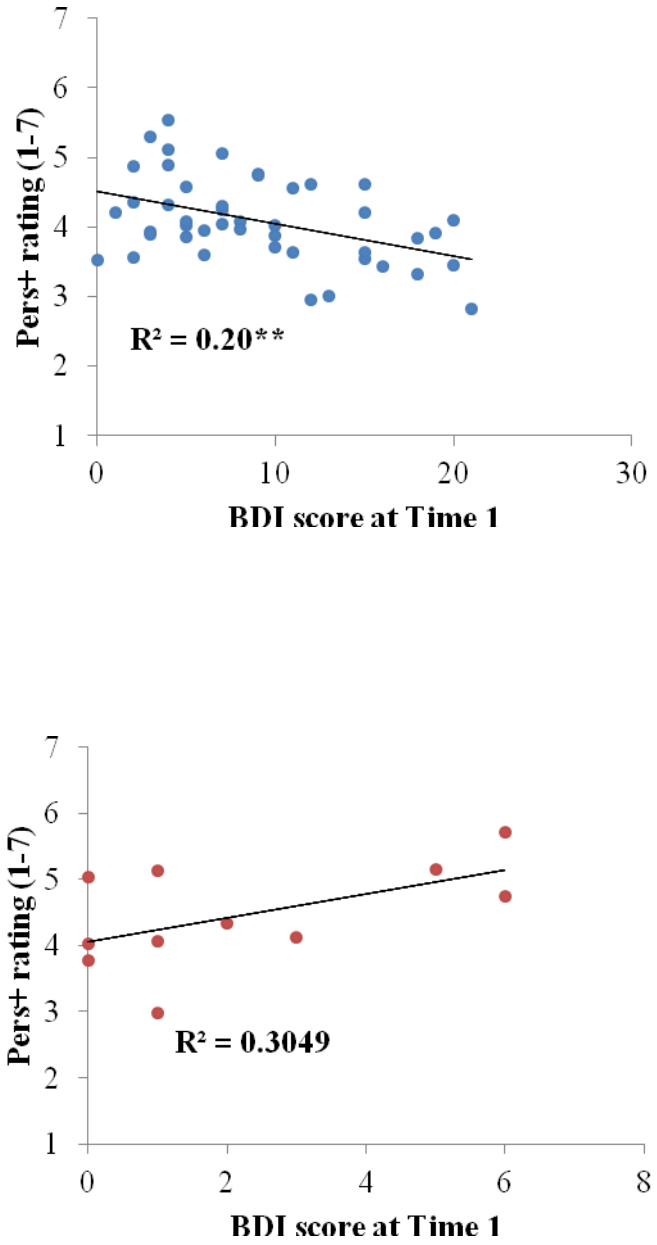


Figure 15. Relationship between BDI scores at Time 1 and Pers+ ratings at Time 1 for PD patients (top) and normal controls (bottom). Higher BDI scores reflect a greater amount of depressive symptomatology. Double asterisk indicates that a relationship was significant at the $p < .01$ level.

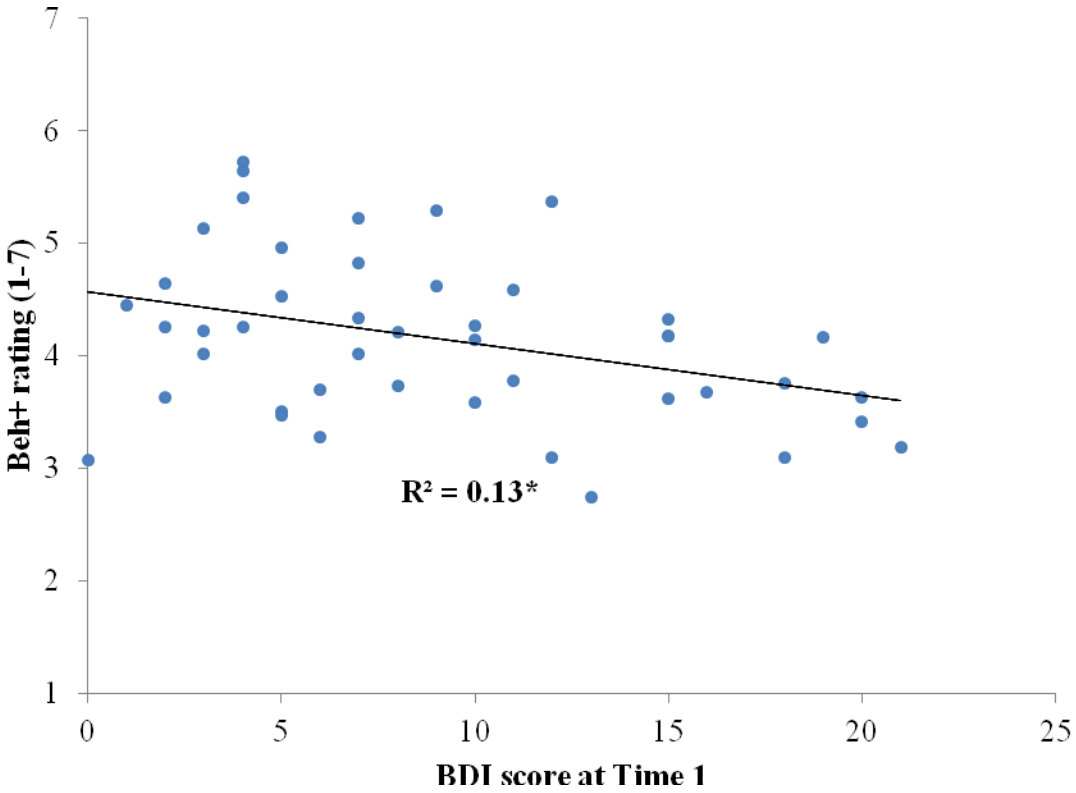


Figure 16a. Relationship between BDI scores at Time 1 and Beh+ ratings at Time 1 for PD patients. Higher BDI scores reflect a greater amount of depressive symptomatology. Asterisk indicates that a relationship was significant at the $p < .05$ level.

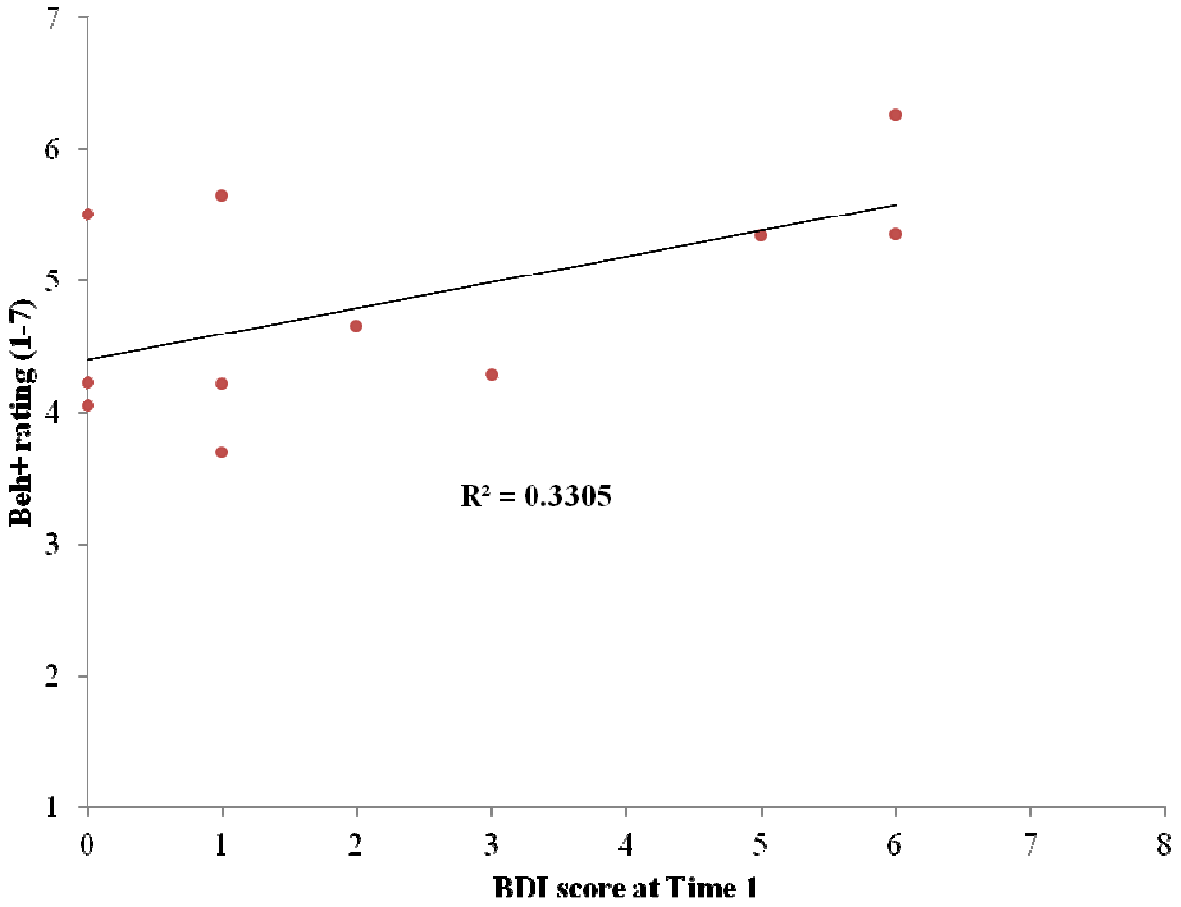


Figure 16b. Relationship between BDI scores at Time 1 and Beh+ ratings at Time 1 for normal controls. Higher BDI scores reflect a greater amount of depressive symptomatology. Asterisk indicates that a relationship was significant at the $p < .05$ level.

Appendix

Neurocognitive, Psychological, and Quality of Life Measures Administered to PD Patients

This test battery consisted of the following assessments: (a) neurocognitive assessments- Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), California Verbal Learning Test – II (Delis, Kramer, Kaplan, & Ober, 2000), Rotary Pursuit Task (Lafayette Instrument Company), Mirror Tracing Task (Lafayette Instrument Company), Judgment of Line Orientation (Benton, Varney & Hamsher, 1975), Boston Naming Test-Short Form (Lansing, Ivnik, Cullum & Randolph, 1999), Delis-Kaplan Executive Function Scales (Sorting, Color-Word Interference, and Verbal Fluency Test subtests only; Delis, Kaplan & Kramer, 2001), American New Adult Reading Test (Grober & Sliwinski, 1991), Self-Ordered Pointing Test of the Colorado Assessment Test (Davis & Keller, 1997), Nine-Hole Peg Test (Mathiowetz, Volland, Kashman, & Weber, 1985), Short Memory subtest of the Repeatable Battery for the Assessment of Neuropsychological Status (Randolph, 1998), and Digit Span subtest of the Wechsler Adult Intelligence Scale III (Wechsler, 1997); (b) psychological assessments- Beck Depression Inventory (2nd version; Beck, Steer, Ball, & Ranieri, 1996), Beck Anxiety Inventory (Beck & Steer, 1990); and (c) quality of life assessments- Parkinson's Disease Questionnaire (US PDQ-39; Bushnell & Martin, 1999), Parkinson's Disease Quality of Life Scale (PDQL; DeBoer, Wijker, Speelman, & de Haes, 1996), Parkinson Disease Activities of Daily Living Scale (PADLS; Hobson, Edwards & Meara, 2001), and EQ-5D (Schrag, Selia, Jahanshahi & Quinn, 2000).

Behavior Rating Form

Video Segment # _ _ _ Your ID: _____

Engaged----------Bored

|----------|----------|----------|----------|----------|----------|

Very engaged	Moderately engaged	Slightly engaged	Neither engaged nor bored	Slightly bored	Moderately bored	Very bored

Tense----------Relaxed

|----------|----------|----------|----------|----------|----------|

Very tense	Moderately tense	Slightly tense	Neither tense nor relaxed	Slightly relaxed	Moderately relaxed	Very relaxed

Behavior Rating Form

Video Segment # _ _ _ Your ID: _____

Sad----------Happy

|----------|----------|----------|----------|----------|----------|

Very sad	Moderately sad	Slightly sad	Neither sad nor happy	Slightly happy	Moderately happy	Very happy

Openhearted----------Guarded

someone who openly revealed his/her experience

someone who was cautious about revealing his/her experience

|----------|----------|----------|----------|----------|----------|

Very openhearted	Moderately openhearted	Slightly openhearted	Neither openhearted nor guarded	Slightly guarded	Moderately guarded	Very guarded

Behavior Rating Form

Video Segment # _ _ _ Your ID: _____

Disliked----------Liked

|----------|----------|----------|----------|----------|----------|----------|

Strongly disliked the person	Moderately disliked the person	Slightly disliked the person	Neither disliked nor liked the person	Slightly liked the person	Moderately liked the person	Strongly liked the person

Personality Rating Form

Video Segment # _ _ _ Your ID: _____

Introverted----------Extraverted

someone who is a reserved private person, does not seek attention, and can be quiet and withdrawn

someone who is outgoing, is talkative, feels comfortable around people, and can be loud and attention-seeking

|----------|----------|----------|----------|----------|----------|----------|

Very introverted | Moderately introverted | Slightly introverted | Neither introverted nor extraverted | Slightly extraverted | Moderately extraverted | Very extraverted

Independent----------Dependent

someone who doesn't care what others think, believes in their own judgment, and feels OK if some people don't like him/her

someone who cares what others think, follows directions, and wants to be liked

|----------|----------|----------|----------|----------|----------|----------|

Very independent | Moderately independent | Slightly independent | Neither independent nor dependent | Slightly dependent | Moderately dependent | Very dependent

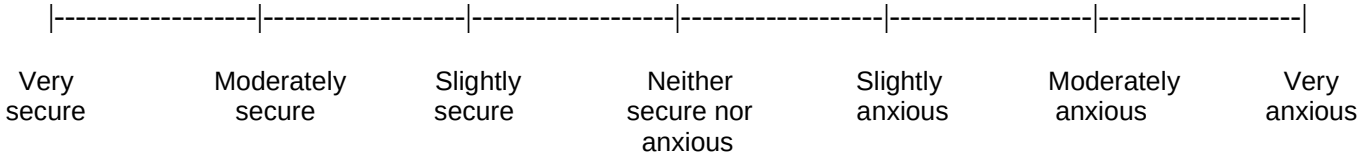
Personality Rating Form

Video Segment # _ _ _ Your ID: _____

Secure----------Anxious

someone who is stable, is calm, feels comfortable with him/herself, and seldom feels blue

someone who is moody, worries, may dislike him/herself, and can often feel blue



Personality Rating Form

Video Segment # _ _ _ Your ID: _____

Disliked----------Liked

|----------|----------|----------|----------|----------|----------|----------|

Strongly disliked the person	Moderately disliked the person	Slightly disliked the person	Neither disliked nor liked the person	Slightly liked the person	Moderately liked the person	Strongly liked the person

Instructions for Rating Participant Behavior

You will now view 30 video clips of 30-second duration each. In each clip, you will see a person talking about a happy experience he or she had. I will ask you to rate a person in each clip on the 5 scales appearing on your Rating Form. The first 4 scales ask you to rate how a person behaves and appears as he or she produces a monologue: whether they are engaged in the monologue or bored, tense or relaxed, sad or happy, and openhearted or guarded. To clarify what is meant by "openhearted" or "guarded," these terms are defined in the boxes on page 2 (point to the boxes). Please use these definitions when giving a rating. I'll ask you to read these definitions right now and tell me if you have questions about these descriptions.

(Wait for them). Do you have any questions?

Finally, the scale on the last page asks you to rate how much you liked or disliked the person. Please make sure not to forget to rate a person in each clip on that scale.

So, for example, if you think that a person is slightly bored during the monologue, you should circle the place on the scale where it says "slightly bored" (point) and then move on to the next scale.

Importantly, because the clips will be played without any sound, please base all your ratings on each person's facial communication. In other words, your ratings of their behavior should be based on their facial movements and expressions as they relate their happy experiences. Because you will be making judgments based on such limited information, it is very understandable if you use intuition, rather than rational analysis, when giving a rating. If you are not sure about which rating to give, make your best guess and do not skip any responses. Also, please wait for a clip to end before giving your ratings.

Now, let me tell you about the Rating Forms. Because your Rating Forms are numbered, it is important that you use the form on the top for the first clip, then use the form directly underneath it for the second clip, and so on. In other words, please proceed from the top to bottom of the stack without skipping any forms. As soon as you are done with the form, put it on the desk next to you. That will signal to me that you are ready for the next clip.

Because it is important that each of you works independently, I will ask you not to make any comments for the entire duration of the study session. If you have a question, please raise your hand.

If they have cell phones on the desk: I also request that you silence and put away your cell phones, please. Don't start until they have done so.

Are there any questions before we start?

Instructions for Rating Participant Personality

You will now view 30 video clips of 30-second duration each. In each clip, you will see a person talking about a happy experience they had. I will ask you to rate a person in each clip on the 4 scales appearing on your Rating Form. The first 3 scales ask you to rate a person's personality on 3 dimensions: introverted vs. extraverted, independent vs. dependent, and secure vs. anxious. Descriptions of what different personality traits mean appear in boxes on the form. Please read these descriptions right now and tell me if you have any questions.

(Wait for them). Do you have any questions?

Finally, the last scale on the form asks you to rate how much you liked or disliked the person. Please make sure not to forget to rate a person in each clip on that scale.

So, for example, if you think that a person is moderately introverted, you should circle the place on the scale where it says "moderately introverted" (point) and then move on to the next scale.

Importantly, because the clips will be played without any sound, please base your ratings on all the scales on each person's facial communication. In other words, your ratings of their personality should be based on their facial movements and expressions as they relate their happy experiences. Because you will be making judgments based on such limited information, it is very understandable if you use intuition, rather than rational analysis, when giving a rating. If you are not sure about which rating to give, make your best guess. Do not skip any responses. Also, please wait for a clip to end before giving your ratings.

Now, let me tell you about how to use the Rating Forms. Because your Rating Forms are numbered, it is important that you use the form on the top for the first clip, then use the form directly underneath it for the second clip, and so on. In other words, please proceed from the top to bottom of the stack without skipping any forms. As soon as you are done with the form, put it on the desk next to you. That will signal to me that you are ready for the next clip.

Because it is important that each of you works independently, I will ask you not to make any comments for the entire duration of the study session. If you have a question, please raise your hand.

If they have cell phones on the desk: I also request that you silence and put away your cell phones, please. Don't start until they have done so.

Are there any questions before we start?

Demographics and Medical History

Rater ID #:

What is your gender:

Which of the following census categories best describes your race/ethnicity:

American Indian or Alaskan Native

Asian

Black or African American

Hispanic, Latino, or Chicano

Indian

Middle Eastern

Native American

Native Hawaiian or Pacific Islander

White

Multiracial (please specify):

Other (please specify):

What is your age:

What is the highest educational degree attained:

How many years have you spent in college so far:

Do you have a history of…

1) …head injury Y/N

2) …neurological disorder Y/N

3) …learning disability Y/N

4) …psychiatric disorder Y/N

5) …alcohol/substance abuse Y/N

Bibliography

Aarsland, D., Andersen, K., Larsen, J. P., Lolk, A., & Kragh-Sorensen, P. (2003). Prevalence and characteristics of dementia in Parkinson disease: an 8-year prospective study. Archives of Neurology, 60, 387-392.

Albright, L., Malloy, T. E., Dong, Q., Kenny, D. A., Fang, X., Winquist, L., & Yu, D. (1997). Cross-cultural consensus in personality judgments. Journal of Personality and Social Psychology, 72, 558-569.

Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. Trends in Neurosciences, 13, 266-271.

Alexander, G. E., Crutcher, M. D., & DeLong, M. R. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. Progress in Brain Research, 85, 119-146.

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual Review of Neuroscience, 9, 357-381.

Allan, C. M. (1970). Treatment of non fluent speech resulting from neurological disease -- treatment of dysarthria. British Journal of Disorders of Communication, 5, 3-5.

Almeida, O. P., & Almeida, S. A. (1999). Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. International Journal of Geriatric Psychiatry, 14, 858-865.

Aparicio, P., Diedrichsen, J., & Ivry, R. B. (2005). Effects of focal basal ganglia lesions on timing and force control. Brain and Cognition, 58, 62-74.

Arabia, G., Grossardt, B. R., Colligan, R. C., Bower, J. H., Maraganore, D. M., Ahlskog, J.

E.,…Rocca, W. A. (2010). Novelty seeking and introversion do not predict the long-term risk of Parkinson disease. Movement Disorders, 25, 2105-2113.

Archer, T., Fredriksson, A., & Johansson, B. (2011). Exercise alleviates Parkinsonism: clinical and laboratory evidence. Acta Neurologica Scandinavica, 123, 73-84.

Bakeman, R., Quera, V., & Gnisci, A. (2009). Observer Agreement for Timed-Event Sequential Data: A Comparison of Time-Based and Event-Based Algorithms. Behavior Research Methods, 41, 137-147.

Bard, C., Turrell, Y., Fleury, M., Teasdale, N., Lamarre, Y., & Martin, O. (1999). Deafferentation and pointing with visual double-step perturbations. Experimental Brain Research, 125, 410-416.

Beck, A. T. (1967). Depression: Clinical, experimental and theoretical aspects. New York: Harper & Row.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). Anxiety disorders and phobias: A cognitive perspective. New York: Basic Books.

Beck, A. T., & Steer, R. A. (1990). Beck Anxiety Inventory (BAI). San Antonio, TX: The Psychological Corporaton.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of the Beck Depression Inventories-IA and -II in psychiatric outpatients. Journal of Personality Assessment, 67, 588-797.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Manual for Beck Depression Inventory II (BDI-II). San Antonio, TX: Psychology Corporation.

Benton, A. L., Varney, N. R., & Hamsher, K. de S. (1975). Judgment of line orientation. Iowa City: University of Iowa, Department of Neurology.

Berardelli, A., Rothwell, J. C., Thompson, P. D., & Hallett, M. (2001). Pathophysiology of bradykinesia in Parkinson's disease. Brain, 124, 2131–2146.

Best, C. H., & Taylor, N. B. (1966). The physiological basis of medical practice (8th ed.). Baltimore: Williams & Wilkins.

Borkenau, P. & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. Journal of Personality and Social Psychology, 62, 645-657.

Borod, J. C., Alpert, M., Brozgold, A., & Martin, C. (1989). A preliminary comparison of flat affect schizophrenics and brain-damaged patients on measures of affective processing. Journal of Communication Disorders, 22, 93-104.

Borod, J. C., Tabert, M. H., Santschi, C., & Strauss, E. (2000). Neuropsychological assessment of emotional processing in brain-damaged patients. In J. C. Borod (Ed.), The neuropsychology of emotion (pp. 80-105). New York: Oxford University Press.

Borod, J. C., Welkowitz, J., Alpert, M., Brozgold, A. Z., Martin, C., Peselow, E., & Diller, L. (1990). Parameters of emotional processing in neuropsychiatric disorders: conceptual issues and a battery of tests. Journal of Communication Disorders, 23, 247-271.

Borod, J. C., Welkowitz, J., & Obler, L. K. (1992). The New York Emotion Battery. Unpublished materials, Mount Sinai Medical Center, Department of Neurology, New York.

Bourne, V. J. (2008). Examining the relationship between degree of handedness and degree of cerebral lateralization for processing facial emotion. Neuropsychology, 22, 350-356.

Bower, J. H., Grossardt, B. R., Maraganore, D. M., Ahlskog, J. E., Colligan, R. C., Geda, Y. E., …Rocca, W. A. (2010). Anxious personality predicts an increased risk of Parkinson's disease. Movement Disorders, 25, 2105-2113.

Bowers, D., Miller, K., Bosch, W., Gokcay, D., Pedraza, O., Springer, U., & Okun, M. (2006a). Faces of emotion in Parkinson's disease: micro-expressivity and bradykinesia during voluntary facial expressions. Journal of the International Neuropsychological Society, 12, 765-773.

Bowers, D., Miller, K., Mikos, A., Kirsch-Darrow, L., Springer, U., Fernandez, H.,…Okun, M. (2006b). Startling facts about emotion in Parkinson's disease: blunted reactivity to aversive stimuli. Brain, 129, 3356-3365.

Brod, M., Mendelsohn, G. A., & Roberts, B. (1998). Patient's experiences of Parkinson's disease. Journal of Gerontology, 53B, 213–222.

Brotchie, P., Iansek, R., & Horne, M. K. (1991). Motor function of the monkey globus pallidus. 2. Cognitive aspects of movement and phasic neuronal activity. Brain, 114, 1685-1702.

Brozgold, A. Z., Borod, J. C., Martin, C. C., Pick, L. H., Alpert, M., & Welkowitz, J. (1998). Social functioning and facial emotional expression in neurological and psychiatric disorders. Applied Neuropsychology, 5, 15-23.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. Psychological Review, 62, 193-217.

Buck, R., & Duffy, R. J. (1980). Nonverbal communication of affect in brain-damaged patients. Cortex, 16, 351-362.

Burn, D. J. (2002). Beyond the iron mask: Towards better recognition and treatment of depression associated with Parkinson's disease. Movement Disorders, 17, 445-454.

Bushnell, D. M., & Martin, M. L. (1999). Quality of life and Parkinson's disease: translation and validation of the US Parkinson's Disease Questionnaire (PDQ-39). Quality of Life Research, 8, 345-50.

Cakit, B. D., Saracoglu, M., Genc, H., Erdem, H. R., & Inan, L. (2007). The effects of incremental speed-dependent treadmill training on postural instability and fear of falling in Parkinson's disease. Clinical Rehabilitation, 21, 698–705.

Camp, C. D. (1913). Paralysis agitans, multiple sclerosis and their treatment. In W. A. White, S. E. Jelliffe, & H. Kimpton (Eds.), Modern treatment of nervous and mental diseases (Vol. 1, pp. 651–671). Philadelphia: Lea & Febiger.

Camras, L. A., Oster, H., Campos, J., Campos, R., Ujiie, T., Miyake, K., & Meng, Z. (1998). Production of emotional facial expressions in European American, Japanese, and Chinese infants. Developmental Psychology, 34, 616-628.

Clark, L. A., & Watson, D. (1991a). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. Journal of Abnormal Psychology, 100, 316-336.

Clark, L. A., & Watson, D. (1991b). Theoretical and empirical issues in differentiating depression from anxiety. In J. Becker, & A. Kleinman (Eds.), Psychosocial aspects of depression (pp. 39-65). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., Cohen, P., West, S. J. & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cohn, J. F. & Schmidt, K. (2003). The timing of facial motion in posed and spontaneous smiles. In J. P. Li, J. Zhao, J. Liu, & J. Yen (Eds.), Proceedings of the Second International Conference on Active Media Technology (pp. 57-69). Hackensack, NJ: World Scientific Publishing Company.

Cole, D. A., Maxwell, S. E., Arvey, R. & Salas, E. (1994). How the power of a MANOVA can both increase and decrease as a function of intercorrelations among the dependent

variables. Psychological Bulletin, 115, 465-474.

Cottrell, C. A., Neuberg, S. L., & Li, N. P. (2007). What do people desire in others? A sociofunctional perspective on the importance of different valued characteristics. Journal of Personality and Social Psychology, 92, 208-231.

Cummings, J. L., & Masterman, D. L. (1999). Depression in patients with Parkinson's disease. International Journal of Geriatric Psychiatry, 14, 711-718.

Dakof, G. A., & Mendelsohn, G. A. (1989). Pattems of adaptation to Parkinson's disease. Health Psychology, 8, 355-372.

Davis, H. P., & Keller, F. R. (1997). Colorado Assessment Tests. Colorado Springs: Colorado Assessment Tests, pp. 1-125.

Deuschl, G. & Goddemeier, C. (1998). Spontaneous and reflex activity of facial muscles in dystonia, Parkinson's disease and in normal subjects. Journal of Neurology, Neurosurgery, and Psychiatry, 64, 320–324.

de Boer, A. G., Wijker, W., Speelman, J. D., & de Haes, J. C. (1996). Quality of life in patients with Parkinson's disease: development of a questionnaire. Journal of Neurology Neurosurgery, and Psychiatry, 61, 70–74.

Delis, D., Kaplan, E., & Kramer, J. (2001). D-KEFS Executive Function System: Examiner's Manual. San Antonio, TX: Psychological Corporation, a Harcourt Assessment Company.

Delis, D., Kramer, J., Kaplan, E., & Ober, B. (2000). CVLT-II California Verbal Learning Test, Second Edition, Adult Version. San Antonio, TX: Psychological Corporation, a Harcourt Assessment Company.

Desmurget, M., Grafton, S. T., Vindras, P., Grea, H., & Turner, R. S. (2003). Basal ganglia network mediates the control of movement amplitude. Experimental Brain Research, 153,

197-209.

Desmurget, M., Grafton, S. T., Vindras, P., Grea, H., & Turner, R. S. (2004). The basal ganglia network mediates the planning of movement amplitude. The European Journal of Neuroscience, 19, 2871-2880.

Drago, V., Foster, P. S., Skidmore, F., Trifiletti, D., & Heilman, K. M. (2008). Spatial emotional akinesia in Parkinson disease. Cognitive and Behavioral Neurology, 21, 92-97.

Dromey, C., & Ramig, L. O. (1998). Intentional changes in sound pressure level and rate: their impact on measures of respiration, phonation, and articulation. Journal of Speech, Language, and Hearing Research, 41, 1003–1018.

Duvoisin, R. C., Eldridge, R., Williams, A., Nutt, J., & Calne, D. (1981). Twin study of Parkinson disease. Neurology, 31, 77– 80.

Ebersbach, G., Ebersbach, A., Edler, D., Kaufhold, O., Kusch, M., Kupsch, A., & Wissel, J. (2010). Comparing exercise in Parkinson's disease—the Berlin BIG Study. Movement Disorders, 25, 1902-1908.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. Semiotica, 1, 49-98.

Ekman, P., & Friesen, W. V. (1978). Facial action coding system: A technique for the measurement of facial movement. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. Journal of Nonverbal Behavior, 6, 238-252.

Ekman, P., Friesen, W. V., Hager, J. C. (2002a). Facial Action Coding System: Investigator’s Guide. Research Nexus, Salt Lake City.

Ekman, P., Friesen, W. V., Hager, J. C. (2002b). Facial Action Coding System. The Manual on

CD ROM. Research Nexus, Salt Lake City.

Ekman, P. & Rosenberg, E. L. (Eds.). (2005). What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS) (2nd ed.). New York: Oxford University Press.

Farley, B., Fox, C., Ramig, L., & McFarland, D. (2008). Intensive amplitude-specific therapeutic approaches for Parkinson disease: toward a neuroplasticity-principled rehabilitation model. Topics in Geriatric Rehabilitation, 24, 99-114.

Fisher, B. E., Wu, A. D., Salem, G. J., Song, J., Lin, C. H., Yip, J.,…Petzinger, G. (2008). The effect of exercise training in improving motor performance and corticomotor excitability in people with early Parkinson's disease. Archives of Physical Medicine and Rehabilitation, 89, 1221-1229.

Fisher, B., Petzinger, G., Nixon, K., Hogg, E., Bremmer, S., Meshul, C., & Jakowec, M. (2004). Exercise-induced behavioural recovery and neuroplasticvity in the 1-methyl-4-phenyl-1,2,3,6-tetrahydroyridine- lesioned mouse basal ganglia. Journal of Neuroscience Research, 77, 378-390.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12, 189-198.

Frank, M. G., Ekman, P., & Friesen, W. V. (1993). Behavioral markers and recognizability of the smile of enjoyment. Journal of Personality and Social Psychology, 64, 83-93.

Frisina, P. G., Borod, J., Foldi, N. S. & Tenenbaum, H. R. (2008). Depression in Parkinson's disease: Health risks, etiology, and treatment options. Neuropsychiatric Disease and Treatment, 4, 81-91.

Fujii, C., Harada, S., Ohkoshi, N., Hayashi, A., & Yoshizawa, K. (2000). Crosscultural traits for personality of patients with Parkinson's disease in Japan. American Journal of Medical Genetics, 96, 1–3.

Gaebel, W. & Wölwer, W. (2004). Facial expressivity in the course of schizophrenia and depression. European Archives of Psychiatry and Clinical Neuroscience, 254, 335-342.

Galvan, A., & Wichmann, T. (2008). Pathophysiology of parkinsonism. Clinical Neurophysiology, 119, 1459-1474.

Gehricke, J.-G. & Shapiro, D. (2000). Reduced facial expression and social context in major depression: Discrepancies between facial muscle activity and self-reported emotion. Psychiatry Research, 95, 157-167.

Georgopoulos, A. P., DeLong, M. R., & Crutcher, M. D. (1983). Relations between parameters of step-tracking movements and single cell discharge in the globus pallidus and subthalamic nucleus of the behaving monkey. Journal of Neuroscience, 3, 1586-1598.

Glatt, S. L., Hubble, J. P., Lyons, K., Paolo, A., Tröster, A. I., Hassanein, R. E. S., & Koller, W. C. (1996). Risk factors for dementia in Parkinson's disease: Effect of education. Neuroepidemiology, 15, 20-25.

Gosselin, P., Perron, M., Legault, M., & Campanella, M. (2002). Children's and adults' knowledge of the distinction between enjoyment and nonenjoyment Smiles. Journal of Nonverbal Behavior, 26, 83-108.

Gracco, V., & Löfqvist, A. (1994). Speech motor coordination and control: Evidence from lip, jaw and laryngeal movements. Journal of Neuroscience, 14, 6585–6597.

Grober, E., & Sliwinski, M. (1991). Development and validation of a model for estimating premorbid verbal intelligence in the elderly. Journal of Clinical and Experimental

Neuropsychology, 13, 933–949.

Gross, J. J., & John, O. P. (1995). Facets of emotional expressivity: Three self-report factors and their correlates. Personality and Individual Differences, 19, 555–568.

Gross, J. J., & John, O. P. (1997). Revealing feelings: Facets of emotional expressivity in self-reports, peer ratings, and behavior. Journal of Personality and Social Psychology, 72, 435– 448.

Gurney, K., Prescott, T. J., & Redgrave, P. (2001). A computational model of action selection in the basal ganglia. I. A new functional anatomy. Biological Cybernetics, 84, 401-410.

Gurney, K., Prescott, T. J., & Redgrave, P. (2001). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. Biological Cybernetics, 84, 411-423.

Halfacre, M., Borod, J., Pick, L., Krch, D. & Gruber, C. (2006). Emotional and nonemotional lexical expression in Parkinson's disease. Paper presented at the annual meeting of the International Neuropsychological Society, Boston, MA.

Harwell, M. R., & Serlin, R. C. (1988). An empirical study of a proposed test of nonparametric analysis of a covariance. Psychological Bulletin, 16, 268-281.

Hemmesch, A. R., Tickle-Degnen, L., & Zebrowitz, L. A. (2009). The influence of facial masking and sex on older adults' impressions of individuals with Parkinson's disease. Psychology and Aging, 24, 542–549.

Herman, T., Giladi, N., Gruendlinger, L., & Hausdorff, J. M. (2007). Six weeks of intensive treadmill training improves gait and quality of life in patients with Parkinson's disease: a pilot study. Archives of Physical Medicine and Rehabilitation, 88, 1154-8.

Hettmansperger, T. P. (1984). Statistical inference based on ranks. New York: Wiley.

Ho, A. K., Bradshaw, J. L., Iansek, R., & Alfredson, R. (1999). Speech volume regulation in Parkinson's disease: effects of implicit cues and explicit instructions. Neuropsychologia, 37, 1453-1460.

Ho, A. K., Iansek, R., & Bradshaw, J. L. (1999). Regulation of parkinsonian speech volume: the effect of interlocuter distance. Journal of Neurology, Neurosurgery, and Psychiatry, 67, 199-202.

Hobson, J. P., Edwards, N. I., & Meara, R. J. (2001). The Parkinson's Disease Activities of Daily Living Scale: a new simple and brief subjective measure of disability in Parkinson's disease. Clinical Rehabilitation, 15, 241–246.

Hoehn, M. M. & Yahr, M. D. (1997). Unified Parkinson's disease rating scale [UPDRS]: Modified Hoehn and Yahr staging. In R. M. Herndon (Ed.), Handbook of clinical neurologic scales. New York: Dernos Vermande.

Hopf, H. C., Muller-Forell, W., & Hopf, N. J. (1992). Localization of emotional and volitional facial paresis. Neurology, 42, 1918-1923.

International Personality Item Pool. (2010). A scientific collaboratory for the development of advanced measures of personality traits and other individual differences. Retrieved April 15, 2010, from Oregon Research Institute website: http://ipip.ori.org

Jacobs, D. H., Shuren, J., Bowers, D., & Heilman, K. M. (1995). Emotional facial imagery, perception, and expression in Parkinson's disease. Neurology, 45, 1696-1702.

Jaeger, J., Borod, J. C., & Peselow, E. (1986). Facial expression of positive and negative emotions in patients with unipolar depression. Journal of Affective Disorders, 11, 43-50.

Jenkinson, C., Peto, V., Fitzpatrick, R., Greenhall, R., & Hyman, N. (1995). Self-reported functioning and well being in patients with Parkinson's disease: Comparison of the Short

Form Health Survey (SF-36) and the Parkinson's Disease Questionnaire (PDQ-39). Age and Ageing, 24, 505-509.

Jiang, J., Alwan, A., Bernstein, L. E., Keating, P., & Auer, E. T. (2000). On the correlation between facial movements, tongue movements, and speech acoustics. In Proceedings from the International Congress on Spoken Language Processing: Vol. 1 (pp. 42-45). Beijing, China.

Joel, D., & Weiner, I. (1994). The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. Neuroscience, 63, 363-379.

Jox, R., Bruning, R., Hamann, G., & Danek, A. (2004). Volitional facial palsy after a vascular lesion of the supplementary motor area. Neurology, 63, 756-757.

Jueptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S., & Passingham, R. E. (1997). Anatomy of Motor Learning. I. Frontal Cortex and Attention to Action. Journal of Neurophysiology, 77, 1313-1324.

Kaasinen, V., Nurmi, E., Bergman, J., Eskola, O., Solin, O., Sonninen, P., & Rinne, J. O. (2001). Personality traits and brain dopaminergic function in Parkinson’s disease. Proceedings of the National Academy of Sciences, USA, 98, 13272–13277.

Karson, C. N. (1983). Spontaneous eye-blink rates and dopaminergic systems. Brain, 106, 643–653.

Katsikitis, M., & Pilowsky, I. (1988). A study of facial expression in Parkinson's disease using a novel microcomputer-based method. Journal of Neurology, Neurosurgery, and Psychiatry, 51, 362-366.

Katsikitis, M., & Pilowsky, I. (1991). A controlled quantitative study of facial expression in Parkinson's disease and depression. The Journal of Nervous and Mental disease, 179,

683-688.

Kazandjian, S., Borod, J. C., & Brickman, A. M. (2007). Facial expression during emotional monologues in unilateral stroke: an analysis of monologue segments. Applied Neuropsychology, 14, 235-246.

Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: A review. British Journal of Mathematical and Statistical Psychology, 54, 1-20.

Keus, S. H., Bloem, B. R., Hendriks, E. J., Bredero-Cohen, A. B., & Munneke, M. (2007). Evidence-based analysis of physical therapy in Parkinson's disease with recommendations for practice and research. Movement Disorders, 22, 451–460.

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. Journal of Nonverbal Behavior, 20, 165-182.

Kuzis, G., Sabe, L., Tiberti, C., Leiguarda, R., & Starkstein, S. (1997). Cognitive functions in major depression and Parkinson's disease. Archives of Neurology, 54, 982–986.

Lansing, A. E., Ivnik, R. J., Cullum, C. M., & Randolph, C. (1999). An empirically derived short form of the Boston Naming Test. Archives of Clinical Neuropsychology, 1, 481–487.

Larrieu, S., Letenneur, L., Orgogozo, J. M., Fabrigoule, C., Amieva, H., Le Carret, N.,…Dartiques, J. F. (2002). Incidence and outcome of mild cognitive impairment in a population-based
prospective cohort. Neurology, 59, 1594-1599.

Lee, Y., & Ottati, V. (1993). Determinants of in-group and out-group perceptions of heterogeneity: An investigation of Sino-American stereotypes. Journal of Cross-Cultural Psychology, 24, 298-318.

Leentjens, A. F. (2004). Depression in Parkinson's disease: conceptual issues and clinical challenges. Journal of Geriatric Psychiatry and Neurology, 17,1200–1206.

Lehéricy, S., Benali, H., Van de Moortele, P. F., Pélégrini-Issac, M., Waechter, T., Ugurbil, K., & Doyon, J. (2005). Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. Proceedings of the National Academy of Sciences, USA, 102, 12566–12571.

Lieberman, A. (2006). Depression in Parkinson's disease – a review. Acta Neurologica Scandinavica, 113,1-8.

Liotti, M., Ramig, L. O., Vogel, D., New, P., Cook, C. I., Ingham, R. J.,…Fox, P. T. (2003). Hypophonia in Parkinson's disease: neural correlates of voice treatment revealed by PET. Neurology, 60, 432-440.

Lix, L. M., & Keselman, H. J. (1995). Approximate degrees of freedom tests: A unified perspective on testing for mean equality. Psychological Bulletin, 117, 547-560.

Ludlow, C. L., & Bassich, C. J. (1984). Relationships between perceptual ratings and acoustic measures of hypokinetic speech. In M. R. McNeil, J. C. Rosenbek, & A. Aronson (Eds.), The dysarthrias: Physiology-acoustic-perception-management (pp. 163-195). San Diego, CA: College Hill Press.

Lyons, K. D., Tickle-Degnen, L., & DeGroat, E. J. (2005). Inferring personality traits of clients with Parkinson's disease from their descriptions of favourite activities. Clinical Rehabilitation, 19, 799-809.

Lyons, K. D., Tickle-Degnen, L., Henry, A., & Cohn, E. (2004a). Behavioural cues of personality in Parkinson's disease. Disability and Rehabilitation, 26, 463-470.

Lyons, K. D., Tickle-Degnen, L., Henry, A., & Cohn, E. (2004b). Impressions of Personality in

Parkinson's disease: Can rehabilitation practitioners see beyond the symptoms? Rehabilitation Psychology, 49, 328-333.

Madeley, P., Ellis, A. W., & Mindham, R. H. S. (1995). Facial expressions and Parkinson's disease. Behavioural Neurology, 8, 115-119.

Marr, J. (1991). The experience of living with Parkinson's disease. Journal of Neuroscience Nursing, 23, 325-329.

Marsh, L. (2000). Neuropsychiatric aspects of Parkinson's disease. Psychosomatics, 41, 15–23.

Marsh, A. A., Adams, R. B., Jr., & Kleck, R. E. (2005). Why do fear and anger look the way they do? Form and social function in facial expressions. Personality and Social Psychology Bulletin, 31, 73-86.

Mathiowetz, V., Weber, K., Kashman, N., & Volland, G. (1985). Adult norms for 9 Hole Peg Test of finger dexterity. Occupational Therapy Journal of Research, 5, 24–38.

McClean, M. D., & Tasko, S. M. (2002). Association of orofacial with laryngeal and respiratory motor output during speech. Experimental Brain Research, 146, 481–489.

Menza, M. A., Golbe, L. I., Cody, R. A., & Forman, N. E. (1993). Dopamine-related personality traits in Parkinson's disease. Neurology, 43, 505–508.

Menza, M. A., Robertson-Hoffman, D. E., & Bonapace, A. S. (1993). Parkinson's disease and anxiety: comorbidity with depression. Biological Psychiatry, 34, 465–470.

Mikos, A. E., Springer, U. S., Nisenzon, A. N., Kellison, I. L., Fernandez, H. H., Okun, M. S., & Bowers, D. (2009). Awareness of expressivity deficits in non-demented Parkinson disease. Clinical Neuropsychologist, 23, 805-817.

Miller, N., Allcock, L., Jones, D., Noble, E., Hildreth, A. J., & Burn, D. J. (2007). Prevalence and pattern of perceived intelligibility changes in Parkinson's disease. Journal of

Neurology, Neurosurgery & Psychiatry, 78, 1188-1190.

Miller , N., Noble, E., Jones, D., Allcock, L., & Burn, D. (2008). How do I sound to me? Perceived changes in communication in Parkinson's disease . Clinical Rehabilitation, 22, 14 – 22.

Miller, N., Noble, E., Jones, D., & Burn, D. (2006). Life with communication changes in Parkinson's disease. Age and Ageing, 35, 235–239.

Mink, J. W., & Thach, W. T. (1991). Basal ganglia motor control. I. Nonexclusive relation of pallidal discharge to five movement modes. Journal of Neurophysiology, 65, 273-300.

Mohr, B., Müller, V., Mattes, R., Rosin, R., Federmann, B., Strehl, U.,…Birbaumer, N. (1996). Behavioral treatment of Parkinson's disease leads to improvement of motor skills and tremor reduction. Behavior Therapy, 27, 235–255.

Monrad-Krohn, G. H. (1924). On the dissociation of voluntary and emotional innervation in facial paresis of central origin. Brain: A Journal of Neurology, 47, 22-35.

Monrad-Krohn, G. H. (1957). The third element of speech: prosody in the neuro-psychiatric clinic. The Journal of Mental Science, 103, 326-331.

Montepare, J. M., & Zebrowitz, L. A. (1998). Person perception comes of age: The salience and significance of age in social judgments. In M. P. Zanna (Ed.), Advances in Experimental Social Psychology, Vol. 30. San Diego, CA: Academic Press.

Montreys, C. R., & Borod, J. C. (1998). A preliminary evaluation of emotional experience and expression following unilateral brain damage. International Journal of Neuroscience, 96, 269–283.

Morecraft, R. J., Stilwell-Morecraft, K. S., & Rossing, W. R. (2004). The motor cortex and facial expression: new insights from neuroscience. The Neurologist, 10, 235-249.

Murphy, N. A., Hall, J. A., & Colvin, C. R. (2003). Accurate intelligence assessments in social interaction: Mediators and gender effects. Journal of Personality, 71, 465–493.

Mutch, W. J., Strudwick, A., Roy, S. K., & Downie, A. W. (1986). Parkinson's disease: disability, review, and management. British Medical Journal (Clinical research ed.), 293, 675-677.

Overall, J. E., & Atlas, R. S. (1999). Power of univariate and multivariate analyses of repeated measurements in controlled clinical trials. Journal of Clinical Psychology, 55, 465–485.

Palmer, K., Wang, H., Bäckman, L., Winblad, B., & Fratiglioni, L. (2002). Differential evolution of cognitive impairment in nondemented older persons: results from the Kungsholmen Project. American Journal of Psychiatry, 159, 436-442.

Pentland, B., Gray, J. M., Riddle, W. J., & Pitcairn, T. K. (1988). The effects of reduced non-verbal communication in Parkinson's disease. British Journal of Disorders of Communication, 23, 31-34.

Pentland, B., Pitcairn, T. K., Gray, J. M., & Riddle, W. J. (1987). The effects of reduced expression in Parkinson's disease on impression formation by health professionals. Clinical Rehabilitation, 1, 307-312.

Petzinger, G. M., Walsh, J. P., Akopian, G., Hogg, E., Abernathy, A., Arevalo, P.,…Jakowec, M. W. (2007). Effects of treadmill exercise on dopaminergic transmission in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine- lesioned mouse model of basal ganglia injury. Journal of Neuroscience, 27, 5291- 5300.

Pfann, K. D., Buchman, A. S., Comella, C. L., & Corcos, D. M. (2001). Control of movement distance in Parkinson's disease. Movement Disorders, 16, 1048-1065.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of

emotion: a meta-analysis of emotion activation studies in PET and fMRI. NeuroImage, 16, 331-348.

Pinto, S., Thobois, S., Costes, N., Le Bars, D., Benabid, A. L., Broussolle, E.,…Gentil, M. (2004). Subthalamic nucleus stimulation and dysarthria in Parkinson's disease: a PET study. Brain, 127, 602-615.

Pitcairn, T. K., Clemie, S., Gray, J. M., & Pentland, B. (1990). Non-verbal cues in the self-presentation of parkinsonian patients. British Journal of Clinical Psychology, 29, 177-184.

Pohl, M., Rockstroh, G., Ruckriem, S., Mrass, G., & Mehrholz, J. (2003). Immediate effects of speed-dependent treadmill training on gait parameters in early Parkinson's disease. Archives of Physical Medicine and Rehabilitation,84, 1760-1766.

Ramig, L. O., Fox, C., & Sapir, S. (2004). Parkinson's disease: speech and voice disorders and their treatment with the Lee Silverman Voice Treatment. Seminars in Speech and Language, 25, 169-180.

Ramig L., Pawlas A. A., & Countryman S. (1995). The Lee Silverman Voice Treatment (LSVT): A practical guide for treating the voice and speech disorders in Parkinson disease. National Center for Voice and Speech, University of Iowa.

Ramig, L., Sapir, S., Countryman, S., Pawlas, A., O'Brien, C., Hoehn, M., & Thompson, L. (2001). Intensive voice treatment (LSVT®) for individuals with Parkinson's disease: A two year follow-up. Journal of Neurology, Neurosurgery and Psychiatry, 71, 493-498.

Ramsey, P. H. (1982). Empirical power of procedures for comparing 2 groups on p variables. Journal of Educational Statistics, 7, 139-156.

Randolph, C. (1998). Repeatable Battery for the Assessment of Neuropsychological Status. San

Antonio, TX: Psychological Corporation.

Raskin, S., Borod, J. & Tweedy, J. (1990). Neuropsychological aspects of Parkinson's Disease. Neuropsychology Review, 1, 185-221.

Rinn, W. E. (1984). The neuropsychology of facial expression: a review of the neurological and psychological mechanisms for producing facial expressions. Psychological Bulletin, 95, 52-77.

Rosenberg, E. L. (2005). Introduction: The study of spontaneous facial expressions in psychology. In P. Ekman, & E. L. Rosenberg (Eds.), What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS) (2nd ed., pp. 3-18). New York: Oxford University Press.

Rosenblum, L. D., Miller, R. M. & Sanchez, K. (2007). Lipread me now, hear me better later: Crossmodal transfer of talker familiarity effects. Psychological Science, 18, 392-396.

Ross, R. T., & Mathiesen, R. (1998). Images in clinical medicine. Volitional and emotional supranuclear facial weakness. The New England Journal of Medicine, 338, 1515.

Rossetti, Y., Desmurget, M., & Prablanc, C. (1995). Vectorial coding of movement: vision, proprioception, or both? Journal of Neurophysiology, 74, 457-463.

Roth, A. J. (1988). Welch tests. In S. Kotz & N. L. Johnson (Eds.), Encyclopedia of statistical sciences, Vol. 9 (pp. 586-589). New York: Wiley.

Rule, N. O., Ambady, N., Adams, R. B. Jr., Ozono, H., Nakashima, S., Yoshikawa, S., & Watabe, M. (2010). Polling the face: Prediction and consensus across cultures. Journal of Personality and Social Psychology, 98, 1-15.

Saku, M., & Ellgring, H. (1992). Emotional reactions to odours in Parkinson's disease A clinical application of ethological methods. Journal of Ethology, 10, 47-52.

Sands, I. (1942). The type of personality susceptible to Parkinson's disease. Journal of the Mount Sinai Hospital, 9, 792-794.

Sapir, S., Ramig, L. O., Fox, C. M. (2011). Intensive voice treatment in Parkinson's disease: Lee Silverman Voice Treatment. Expert Review of Neurotherapeutics, 11, 815-830.

Saulsman, L. M., & Page, A. C. (2004). The five-factor model and personality disorder empirical literature: A meta-analytic review. Clinical Psychology Review, 23, 1055-1085.

Sayette, M. A., Cohn, J. F., Wertz, J. M., Perrott, M. A., & Parrott, D. J. (2001). A psychometric evaluation of the Facial Action Coding System for assessing spontaneous expression. Journal of Nonverbal Behavior, 25, 167–185.

Sayette, M. A., Wertz, J. M., Martin, C. S., Cohn, J. F., Perrott, M. A., & Hobel, J. (2005). Effects of smoking opportunity on cue-elicited urge: A facial coding analysis. In P. Ekman, & E. L. Rosenberg (Eds.), What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS) (2nd ed., pp. 583-602). New York: Oxford University Press.

Scherer, R. C. (1991). Physiology of phonation: A review of basic mechanics. In C. N. Ford & D. M. Bless (Eds.), Phonosurgery: Assessment and surgical management of voice disorders (pp. 77-93). New York: Raven.

Schrag, A., Selia, C., Jahanshahi, M., & Quinn, N. (2000). The EQ-5D - a generic quality of life measure - is a useful instrument to measure quality of life in patients with Parkinson's disease. Journal of Neurology, Neurosurgery, and Psychiatry, 69, 67-73.

Schrag, A., Jahanshahi, M., & Quinn, N. (2001). What contributes to depression in Parkinson's disease? Psychological Medicine, 31, 65–73.

Shrout, P. E. & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New

procedures and recommendations. Psychological Methods, 7, 422-445.

Shulman, L. M., Taback, R. L., & Bean, J. (2001). Co-morbidity of the nonmotor symptoms of Parkinson's disease. Movement Disorders, 16, 507–10.

Simons, G., Ellgring, H., & Pasqualini, M. S. (2003). Disturbance of spontaneous and posed facial expressions in Parkinson's disease. Cognition and Emotion, 17, 759-778.

Simons, G., Pasqualini, M. C., Reddy, V., & Wood, J. (2004). Emotional and nonemotional facial expressions in people with Parkinson's disease. Journal of the International Neuropsychological Society, 10, 521-535.

Symons, F. J., Harper, V., Shinde, S. K., Clary, J., & Bodfish, J. W. (2010). Evaluating a sham-controlled sensory-testing protocol for nonverbal adults with neurodevelopmental disorders: self-injury and gender effects. Journal of Pain, 11, 773–781.

Smith, M. C., Smith, M. K., & Ellgring, H. (1996). Spontaneous and posed facial expression in Parkinson's disease. Journal of the International Neuropsychological Society, 2, 383-391.

Smith, M. E., Ramig, L. O., Dromey, C., Perez, K. S., & Samandari, R. (1995). Intensive voice treatment in Parkinson disease: laryngostroboscopic findings. Journal of Voice, 9, 453-459.

Spielman, J. L., Borod, J. C., & Ramig, L. O. (2003). The effects of intensive voice treatment on facial expressiveness in Parkinson disease: preliminary data. Cognitive and Behavioral Neurology, 16, 177-188.

Spraker, M. B., Yu, H., Corcos, D. M., & Vaillancourt, D. E. (2007). Role of individual basal ganglia nuclei in force amplitude generation. Journal of Neurophysiology, 98, 821-834.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. Journal of the International Neuropsychological Society, 8, 448-460.

Stevens, J. (1996). Applied multivariate statistics for the social sciences (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Streifler, M., & Hofman, S. (1984). Disorders of verbal expression in parkinsonism. Advances in Neurology, 40, 385-393.

Sullivan, H. S. (1953). The interpersonal theory of psychiatry. New York: W. W. Norton.

Suteerawattananon, M., Morris, G. S., Etnyre, B. R., Jankovic, J., & Protas, E. J. (2004). Effects of visual and auditory cues on gait in individuals with Parkinson's disease. Journal of Neurological Sciences, 219, 63– 69.

Takahashi, K., Tickle-Degnen, L., Coster ,W. J., & Latham, N. (2010). Expressive behavior in Parkinson's disease as a function of interview context. American Journal of Occupational Therapy, 64, 484–495.

Tickle-Degnen, L., & Lyons, K. D. (2004). Practitioners' impressions of patients with Parkinson's disease: the social ecology of the expressive mask. Social Science & Medicine, 58, 603-614.

Titze, I. R., & Sundberg, J. (1992). Vocal intensity in speakers and singers. The Journal of the Acoustical Society of America, 91, 2936-2946.

Todes, C. J. & Lees, A. J. (1985). The pre-morbid personality of patients with Parkinson's disease. Journal of Neurology, Neurosurgery, and Psychiatry, 48, 97-100.

Tillerson, J. L., Caudle, W. M., Reveron, M. E., & Miller, G. W. (2002). Detection of behavioral impairments correlated to neurochemical deficits in mice treated with moderate doses of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine. Experimental Neurology, 178, 80–90.

Tillerson, J. L., Caudle, W. M., Reveron, M. E., & Miller, G. W. (2003). Exercise induces behavioral recovery and attenuates neurochemical deficits in rodent models of

Parkinson's disease. Neuroscience,119, 899-911.

Tillerson, J. L., Cohen, A. D., Philhower, J., Miller, G. W., Zigmond, M. J., & Schallert, T. (2001). Forced limb-use effects on the behavioral and neurochemical effects of 6-hydroxydopamine. Journal of Neuroscience, 21, 4427–4435.

Trail, M., Fox, C., Ramig, L. O., Sapir, S., Howard, J., & Lai, E. C. (2005). Speech treatment for Parkinson's disease. NeuroRehabilitation, 20, 205-221.

Vaillancourt, D. E., Mayka, M. A., Thulborn, K. R., & Corcos, D. M. (2004). Subthalamic nucleus and internal globus pallidus scale with the rate of change of force production in humans. NeuroImage, 23, 175-186.

Vaillancourt, D. E., Yu, H., Mayka, M. A., & Corcos, D. M. (2007). Role of the basal ganglia and frontal cortex in selecting and producing internally guided force pulses. NeuroImage, 36, 793-803.

Ward, C. D., Duvoisin, R. C., Ince, S. E., Nutt, J. D., Eldridge, R., Calne, D. B., & Dambrosia, J. (1984). Parkinson's disease in twins. Advances in Neurology, 40, 341–344.

Wechsler, D. (1997). Wechsler Adult Intelligence Scale- III (WAIS-III). San Antonio, TX: Psychological Corp.

Weddell, R. A. (1994). Effects of subcortical lesion site on human emotional behavior. Brain and Cognition, 25, 161-193.

Weintraub, D., Moberg, P. J., Duda, J. E., Katz, I. R., & Stern, M. B. (2003). Recognition and treatment of depression in Parkinson's disease. Journal of Geriatric Psychiatry and Neurology, 16, 178 –183.

Weiss, A., Sutin, A. R., Duberstein, P. R., Friedman, B., Bagby, R. M., & Costa, P. T., Jr. (2009). The personality domains and styles of the five-factor model predict incident

depression in Medicare recipients aged 65 to 100. American Journal of Geriatric Psychiatry, 17, 591–601.

Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. Biometrika, 38, 330-336.

Wheeler, P., Haertel, G. D. & Scriven, M. (1992). Teacher Evaluation Glossary. Retrieved July 13, 2009, from Western Michigan University, The Evaluation Center Web site: http://www.wmich.edu/evalctr/ess/glossary/

Widiger, T. A., & Lynam, D. R. (2002). Psychopathy and the five-factor model of personality. In T. Millon, E. Simonsen, M. Birket-Smith, & R. D. Davis (Eds.), Psychopathy: Antisocial, criminal, and violent behavior (pp. 171-187). New York: Guilford Press.

Wieser, M. J., Muhlberger, A., Alpers, G. W., Macht, M., Ellgring, H., & Pauli, P. (2006). Emotion processing in Parkinson's disease: dissociation between early neuronal processing and explicit ratings. Clinical Neurophysiology, 117, 94-102.

Wilcox, R. R. (1990). Comparing the variances of two dependent groups. Journal of Educational Statistics, 15, 237-247.

Woods, S. A. & Hampson, S. E. (2005). Measuring the Big Five with single items using a bipolar response scale. European Journal of Personality, 19, 373-390.

Woolhead, G., Calnan, M., Dieppe, P., & Tadd, W. (2004). Dignity in older age: what do older people in the United Kingdom think? Age and Ageing, 33, 165–70.

Yehia, H., Rubin, P., & Vatikiotis-Bateson, E. (1998). Quantitative association of vocal-tract and facial behavior. Speech Communication, 26, 23–43.

Yekutiel, M. P., Pinhasof, A., Shahar, G., & Sroka, H. (1991). A clinical trial of the re-education of movement in patients with Parkinson’s disease. Clinical Rehabilitation, 5, 207–214.

Yorkston, K. M., Hakel, M., Beukelman, D. R., & Fager, S. (2007). Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review. Journal of Medical Speech-Language Pathology, 15, xi-xxxvi.

Yorkston, K. M., Spencer, K. A., & Duffy, J. R. (2003). Behavioral management of respiratory/phonatory dysfunction from dysarthria: a systematic review of the evidence. Journal of Medical Speech-Language Pathology, 11, xiii–xxxviii.

Zebrowitz, L. A. (1996). Physical appearance as a basis of stereotyping. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), Stereotypes and stereotyping (pp. 79–120). New York: Guilford.

Zebrowitz, L. A. (1997). Reading faces: Window to the soul? Boulder, CO: Westview Press.

Zgaljardic, D. J., Borod, J. C., Foldi, N. S., & Mattis, P. (2003). A review of the cognitive and behavioral sequelae of Parkinson's disease: relationship to frontostriatal circuitry. Cognitive and Behavioral Neurology, 16, 193-210.